PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 34 March 15, 1948 Number 3

A CYTOCHEMICAL STUDY OF THE FEULGEN NUCLEAL REACTION

By Henry S. Di Stefano

Department of Zoology, Columbia University

Communicated by L. C. Dunn, January 10, 1948

The growing interest in the study of the chemical composition of cellular formed elements has led to wide-spread research on techniques for demonstrating and estimating, *in situ*, the presence and amounts of the various components found by macrochemical analysis. In short, this would involve the use of a quantitive chemical reaction and a means of accurately measuring the end-product of the reaction. Such techniques for measuring nucleoproteins of cell nuclei were recently described by Pollister and Ris.[1] Using these same methods, a study of the nature of the Feulgen nucleal reaction was undertaken.

The Feulgen reaction may be considered as occurring in two steps: first, the splitting of the linkage between the purine bases and the desoxypentose of the nucleic acid by means of mild acid hydrolysis, and second, a chemical reaction between the aldehyde groups, thus formed on the sugar, and the Schiff reagent (decolorized fuchsin) resulting in the synthesis of a new dye compound. It was noticed by Bauer,[2] and confirmed by Hillary,[3] that the intensity of color produced in chromatin of cell nuclei is related to hydrolysis in such a way that as time of hydrolysis is increased there is a more and more intense staining until a maximum point is reached; while beyond this maximum, longer hydrolysis results in a decreased staining. The increased staining could easily be explained by the progressive splitting off of bases, as a consequence of which more and more sugar aldehydes are made available for the reaction, which would lead one to expect a maximum when all the purine-desoxyribose linkages have been broken. But the decrease in color after the maximum intensity of stain is reached is quite unexpected. Bauer offered two possible explanations for the latter phenomenon: first, it might be due to a continued decomposition of that part of the nucleic acid molecule remaining after the purine bases have been split off; or second, it

might be due to the destruction of the protein-nucleic acid linkage allowing the latter to diffuse from the chromatin. Hillary subsequently showed that the decrease in intensity could not be due to a chemical degradation of the aldehyde or color-producing groups in the sugar, because a solution of thymus nucleic acid in a test tube is still capable of reaction to produce color after eight hours of hydrolysis. Furthermore, Hillary hydrolyzed nucleic acid agar blocks from five to thirty minutes, and after testing both block and hydrolyzing medium with the Schiff reagent found that the staining capacity of the latter was roughly the reciprocal of that in the agar blocks. This last experiment suggests that one result of prolonged hydrolysis is to change the nucleic acid into a more readily diffusible form.

In order to reach a clearer understanding of what might be happening to the nucleic acid molecule as time of hydrolysis is increased, Professor Arthur W. Pollister suggested that an attempt be made to follow the fate of the molecule by separate photometric determination of each of the three constituents of the individual mononucleotides. Cartilage cells from the head region of frog tadpoles were selected as especially suitable for these measurements, since in these cells the nuclei are nearly perfect spheres and the cytoplasm contains little ultra-violet absorbing material. The diameter of these nuclei is about seven microns. The cartilage was fixed in Carnoy's acetic alcohol (1:3), and sections were cut at ten microns. Only whole nuclei were measured, and corrections were made for the cytoplasmic absorption and for the non-specific light loss as determined in appropriate blanks. The fate of the nucleic acid bases was followed by means of ultra-violet absorption, that of the desoxypentose by measuring the amount of dye produced by the Feulgen reaction performed in the manner described by Stowell,[4] and that of the phosphoric acid by overstaining with methyl green and removing most of the dye not chemically combined, by prolonged washing in absolute isopropyl alcohol.

Bauer constructed curves based only upon visual estimates of Feulgen dye intensity; and the similar curve of Fig. 1, which is based upon actual photometric measurements of absorption by the dye in nuclei, agrees fully with his semi-quantitative observations. There is a peak at 12 minutes which is followed by a decline until at 24 minutes the extinction value is barely measurable.

The change in concentration of purine and pyrimidine bases throughout hydrolysis is shown by the curve based on the ultra-violet absorption. Feulgen[5] found that mild acid hydrolysis, like that in the nucleal reaction, readily breaks the purine-sugar linkage, while by contrast, the pyrimidine-sugar linkage is broken only after much more drastic treatment. It has been recently shown that most, perhaps all, desoxypentose nucleic acids contain equal amounts of purines and pyrimidines (Mirsky and Pollister[6]); consequently, in the Feulgen reaction the time when the ultra-violet ex-

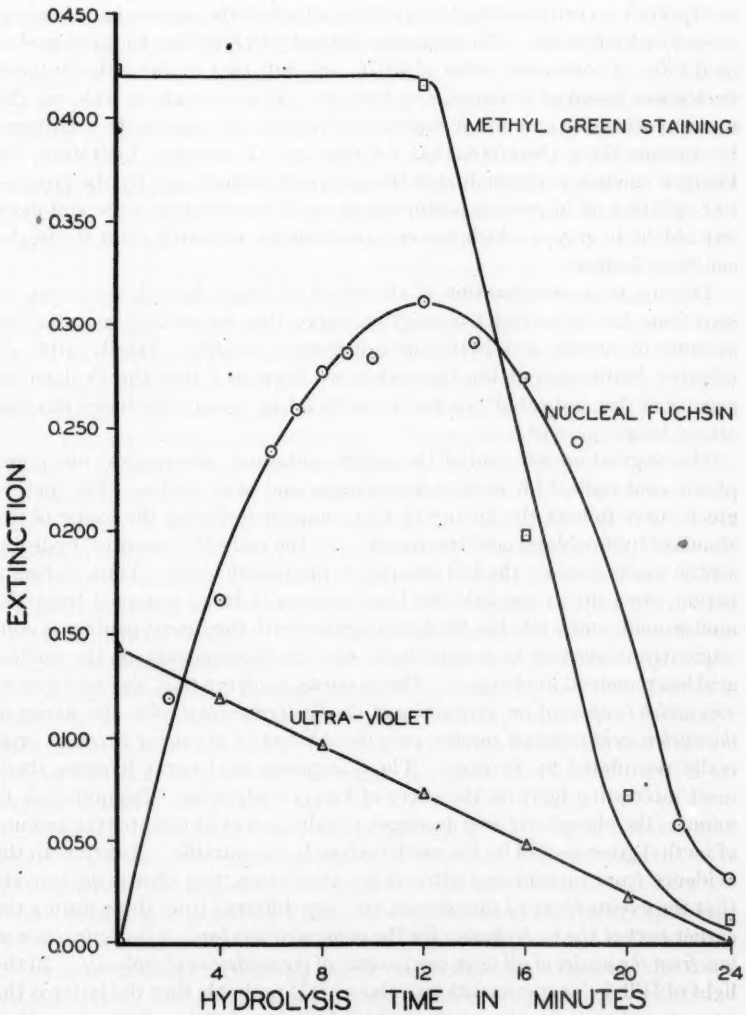

FIGURE 1

Absorption measurements of cartilage cell nuclei from the head region of frog tadpoles at various times of hydrolysis. Each point represents the statistical mean of from 5 to 25 nuclei depending upon the number of nuclei necessary to show a significant difference between consecutive points at a 5% level of significance.

tinction has reached a value one-half that of the unhydrolyzed nuclei would be expected to correspond to that point at which all the purines have been removed by hydrolysis. The extinction for unhydrolyzed nuclei was found to be 0.140. A measured value of 0.070, one-half that of the unhydrolyzed nuclei, was found at 12 minutes hydrolysis. As we saw above, this was the moment when the amount of regenerated fuchsin dye was at its maximum. It appears likely therefore, that for the first 12 minutes' hydrolysis the Feulgen reaction is proceeding in the accepted manner, i.e., by the progressive splitting off of purines producing in the desoxypentose more and more free aldehyde groups which become available for restoring color to the decolorized fuchsin.

Turning to a consideration of the effect of longer hydrolysis, it can be seen from the ultra-violet absorption curve that beyond 12 minutes the amount of purine and pyrimidine decreases steadily. Finally after 24 minutes' hydrolysis, at the time when we have seen that the Feulgen regenerated dye indicated practically no reacting sugar, the bases likewise are no longer present.

The original nucleic acid of the nuclei contained, presumably, one phosphoric acid radical for each desoxypentose and base residue. The methyl green curve follows the history of this component during the course of the standard hydrochloric acid treatment. At the end of 12 minutes' hydrolysis the nuclei contain the full amount of phosphoric acid. Thus, during a period when up to one-half the base content is being removed from the nucleic acid, and while the Feulgen reaction with the desoxypentose is consequently increasing to a maximum, the third component of the nucleic acid has remained unchanged. This is strong evidence that, *during a hydrolysis which is optimal for production of the Feulgen nucleal color, the change in the nucleic acid molecule involves only the splitting off of purine bases,* as originally postulated by Feulgen. The phosphoric acid curve likewise sheds most interesting light on the effect of longer hydrolysis. Beginning at 12 minutes the phosphoric acid decreases rapidly and at 24 minutes the amount of methyl green bound by the nuclei is barely measurable. Taken with the evidence from fuchsin and ultra-violet absorption, this shows conclusively that the events from 12 minutes on are very different from those during the earlier part of the hydrolysis; for the *overhydrolysis leads to the destruction or loss from the nuclei of all three components of the nucleic acid molecule.* In the light of Hillary's experiments (see above) it is probable that the latter is the correct explanation; the overhydrolysis has altered the nucleic acid so that it freely diffuses out of the chromatin. It would seem that this very likely involves some depolymerization. Another process that may be expected to be taking place is a breaking of the nucleic acid-histone salt linkage; for Carnoy's reagent has been found by Pollister and Ris to be unique among cytological fixing fluids in that it preserves the nuclear histones in such a

form that they are readily removable by chemical means (e.g., by solution in the divalent mercury-sulfuric acid Millon reagent).

This brings us to the question of the relation of fixation to the Feulgen reaction, which was exhaustively studied by Bauer. He found sharp-peaked nucleal reaction curves, like that above, after fixation with such fluids as Carnoy and acetic sublimate. A similar curve has also been found for a trichloroacetic sublimate fixing fluid (Pollister unpublished). By contrast, Bauer found that when tissue is fixed in fluids containing chromic acid (such as Allen's Bouin, San Felice, Flemming, Flemming-Heitz, and Champy) instead of a sharply restricted moment of optimal hydrolysis there is a long optimal period during which there is no change in stain intensity; although very prolonged hydrolysis eventually does lead to a drop in stain intensity. It is interesting that the fixing fluids which Bauer found to give sharp-peaked curves upon hydrolysis are those which fix tissue in such a way that all, or nearly all, the histone is readily removable; while from chromic acid fixed material the same treatment removes little or none of the protein (Pollister unpublished). Experiments are now being performed to measure the amount of histone at different stages of hydrolysis.

A compelling reason for undertaking the present study was to examine the possibility of using the Feulgen reaction for quantitative determination of desoxypentose nucleic acid in a cytological preparation. The following considerations indicate that, under proper conditions, the intensity of the Feulgen reaction is an accurate measure of the amount of desoxypentose nucleic acid in a nucleus. The stain intensity at the peak is fairly accurately reproducible for one type of tissue, and the extinction value at the peak after chromic acid fixation seems to be close to that after Carnoy's fluid. As we have seen above, this peak probably represents a moment when all the purines have been removed. From the extinction value of the optimal Feulgen reaction, one can readily compute the amount of dye in an average nucleus (standard value for regenerated fuchsin, 4.4×10^{-6} gram per cc., 1 cm. path, extinction at 530 mμ, 0.745). It is Wieland and Scheuing's[7] view of the nature of the Schiff reaction, that two parts of each decolorized fuchsin molecule react with two aldehyde groups to restore the color. Hence, if it is assumed that the color in the nucleus after optimum hydrolysis has resulted from the reaction of one dye molecule with two desoxypentose residues to which a purine was originally attached, the total amount of nucleic acid in an average nucleus can be computed. A wholly independent computation of nucleic acid in a cartilage cell nucleus may be made from the extinction, near the ultra-violet absorption peak, of nucleic acid in an unhydrolyzed preparation. By calculation from ultra-violet measurements, the average cartilage nucleus contains 2.44×10^{-12} gram of nucleic acid; from the Feulgen reaction the average nucleus contains 2.39×10^{-12} gram of desoxypentose nucleic acid. A certain amount of justi-

fication for the use of methyl green staining as a measure of nucleic acid phosphoric acid comes from the fact that if one assumes that each of the dibasic methyl green molecules unites with the phosphoric acid of two nucleotides the amount of nucleic acid in the average cartilage nucleus is computed to be 2.44×10^{-12} gram.

It should be added that the above figures are of the same order of magnitude as previous determinations of total desoxypentose nucleic acid in a single nucleus. Pollister and Ris report that the nucleus of the calf thymus lymphocyte contains 1.0×10^{-12} gram, from cytological determination by ultra-violet absorption; while calculation from macrochemical data on the lymphocyte nucleus gave practically the same figure, 1.1×10^{-12} gram. Caspersson[8] found 11.2×10^{-12} gram of nucleic acid in the large grasshopper leptotene spermatocyte nucleus, by the ultra-violet cytological method. From the macrochemical data of Zittel and O'Dell[9] the amount of desoxypentose nucleic acid in the nucleus of a single bull spermatozoon can be computed to be 4.6×10^{-12} gram.

Summary.—By photometric measurements the base, sugar and phosphoric acid of the desoxypentose nucleic acid of cartilage nuclei have been followed throughout the standard hydrolysis of the Feulgen nucleal reaction. The results indicate that at optimum hydrolysis time one-half the total base content has been removed; presumably this represents the purines. During post-optimal hydrolysis the whole nucleic acid molecule disappears from the nuclei.

The possibility of using the Feulgen reaction for quantitative determination of desoxypentose nucleic acid is briefly considered.

[1] Pollister, A. W., and Ris, H., *Cold Spring Harbor Symp. Quant. Biol.*, **12**, 147 (1947).
[2] Bauer, H., *Z. Zellforsch. u. mikroskop. Anat.*, **15**, 225 (1932).
[3] Hillary, B. B., *Botan. Gaz.*, **101**, 276 (1939).
[4] Stowell, R. E., *Stain Technol.*, **20**, 45 (1945).
[5] Feulgen, R., *Zeitschr. physiol. Chem.*, **128**, 154 (1923).
[6] Mirsky, A. E., and Pollister, A. W., *J. Gen. Physiol.*, **30**, 117 (1946).
[7] Wieland, H., and Scheuing, *Ber.*, **54**, 2527 (1921).
[8] Caspersson, T., *Chromosoma*, **1**, 147 (1939).
[9] Zittel, C. A., and O'Dell, R. A., *J. Biol. Chem.*, **140**, 899 (1941).

DIETARY FACTORS IN THE UTILIZATION OF HOMOCYSTINE*

BY G. W. KIDDER AND VIRGINIA C. DEWEY

BIOLOGICAL LABORATORY, AMHERST COLLEGE, AMHERST, MASSACHUSETTS

Communicated by E. W. Sinnott, January 15, 1948

Introduction.—It has been known for some time, mainly due to the investigations of du Vigneaud, *et al.*, [1,2] that homocystine can replace methionine in the diet of the rat provided a methyl donor, such as choline or betaine, is present. The recent work of McKittrick[3] has pointed to the same conclusion regarding the chick.

Evidence for the ability of the rat to synthesize methionine from homocystine in the absence of choline, betaine or any other known methyl donor was first presented by Toennies, *et al.*[4] They found that rats, maintained on a rich pre-experimental diet and then on a depletion diet, lacking choline and sulfur containing amino acids, responded to the addition of choline plus homocystine or to homocystine alone. They discuss the influence of intestinal microörganisms on their results and attempt to minimize this influence by the administration of sulfasuxadine[5]. Even after prolonged sulfonamide treatment their rats showed a positive growth response to homocystine but only in the presence of a liver fraction. This liver fraction was essentially free of methionine and choline. The authors concluded that liver contained a factor or factors making possible the methylation of homocystine from sources other than choline or betaine.

In the course of our investigations of the biochemistry of the ciliated protozoan, *Tetrahymena*, we found that homocystine could replace methionine in its diet.[6] Choline was present as was a crude growth factor preparation (Factor II) from liver. When Factor II was prepared from plant sources, homocystine alone did not support growth. In addition, subsequent experiments have shown that exogenous choline does not influence the methylation of homocystine but that our Factor II preparation from liver contained the agent described by Toennies and Bennett, while Factor II prepared from plant material (Cerophyl) was devoid of the Toennies and Bennett factor.

The purpose of this report is to present our findings, which support those of Toennies and Bennett, on the occurrence of a factor or factors in liver which allow the methylation of homocystine to methionine in the absence of exogenous choline and betaine. The possibility of unknown microörganisms vitiating the results, always present in work done with vertebrates, is obviated in our work. *Tetrahymena* is grown in pure (bacteria-free) culture and our results reflect the animal's metabolism only. Moreover, it should be remembered that *Tetrahymena* possesses a rather typical

animal biochemical pattern so that comparisons of our results with those on other animals are permissible.

Experimental.—The organism used in this investiagtion was the ciliated protozoan, *Tetrahymena geleii* W grown in pure (bacteria-free) culture.

In our previous studies on the biochemistry of this organism [7–12] growth responses were measured by counting the number of ciliates, after appropriate dilutions, in measured volumes of media. The results were expressed as organisms per ml. of fluid. Various attempts to measure growth turbidimetrically had failed, due largely to the tendency of these organisms to stratify before readings could be taken. The Lumetron Colorimeter (Model 400-A) which is of the jewel-bearing type and has an extremely rapid response has proved ideal for this type of work.

TABLE 1

BASE MEDIUM (MINUS METHIONINE)

	γ/ML.		γ/ML.
L-arginine	125	Biotin (free acid)	0.0005
L-histidine	125	Ca pantothenate	0.10
DL-isoleucine	125	Thiamine HCl	1.00
L-leucine	250	Nicotinamide	0.10
L-lysine	250	Pyridoxine HCl	0.10
DL-phenylalanine	350	Riboflavin	0.10
DL-threonine	125	Pteroylglutamic acid	0.01
L-tryptophane	50	Choline Cl	1.00
DL-valine	125	Hydrolyzed yeast nucleic acid	100.00
DL-serine	250	Factor II (from Cerophyl) 1:5	.....
		Dextrose	1000.00
		$M_9SO_4.7H_2O$	100.00
		$MgSO_4$	
		K_2HPO_4	100.00
		$CaCl_2.2H_2O$	50.00
		$FeCl_3.6H_2O$	1.25
		$MnCl_2.4H_2O$	0.05
		$ZnCl_2$	0.05

All organisms in experimental series were grown in 4-ml. volumes in 14.5 × 125-mm. Pryrex tubes. The tubes were incubated in a slanted position. The increased volume (2-ml. volumes were previously used) was necessary for use in the colorimeter and the slanted position increased the surface area for oxygenation, thereby increasing the yields. All experiments were carried through three serial transplants, readings being taken only on the third transplant. Our method has been to take readings on all third transplant tubes before inoculation, using distilled water as a blank. All readings are made using red filter 650. After inoculation and incubation a second reading is taken. The optical density of the uninoculated tube is subtracted from the optical density obtained after growth has occurred. This procedure is necessary, since the media used

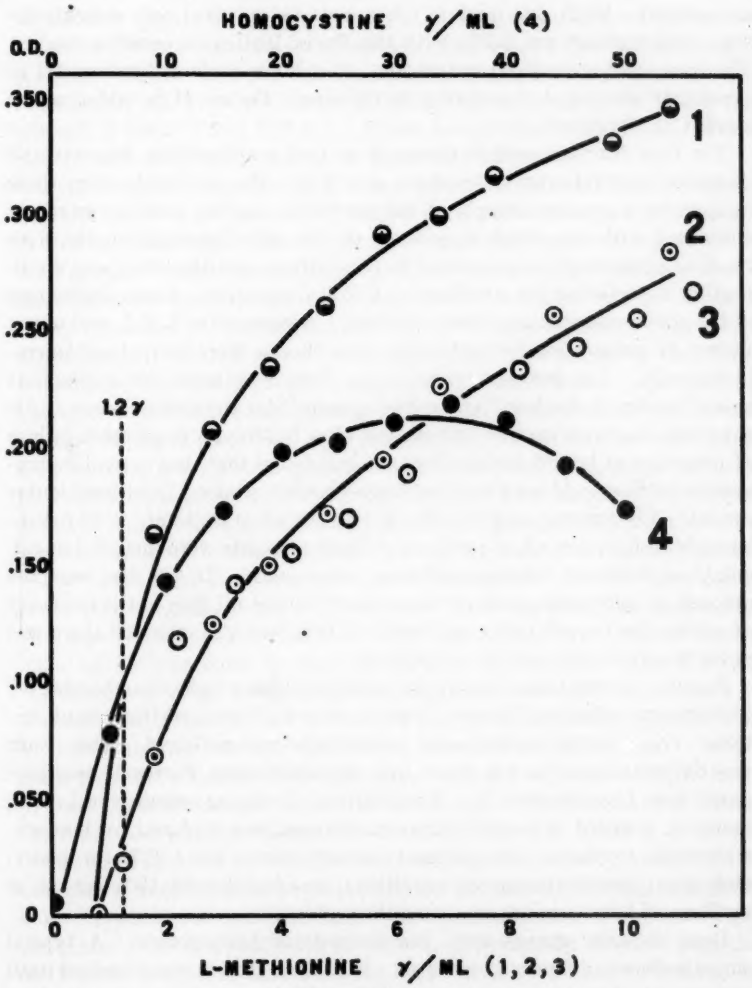

FIGURE 1

Dose response curves. Growth measured in optical density is plotted against amounts in γ per ml. of medium. Curve 1 represents the response to 40γ of homocystine per ml. and graded amounts of methionine. Curve 2 represents response to methionine only. Curve 3 represents response to methionine in the presence of 120γ of L.E.L. Curve 4 represents response to homocystine (upper scale) in the presence of 120γ of L.E.L. (no methionine added). All values from third serial transplants after 6 days incubation.

are colored. With this method tubes need be matched only roughly for size. All readings are taken with the Pyrex label as a position marker. The base medium is given in table 1. This is the same as Medium III as previously described,[12] omitting methionine. Factor II is added as an extract of Cerophyl[8].

The liver fraction used is the same as that employed by Bennett and Toennies[5] and referred to by them as L.E.L. We are indebted to these authors for a generous supply of this material. In this work we were not concerned with the absolute potency of the active material in the liver fraction, accordingly it was added to the medium and the whole was sterilized by autoclaving for 10 minutes at 15 lbs. pressure. Some destruction of the active material may have occurred. Assays of the L.E.L. and of our Factor II preparation for methionine and choline were carried out microbiologically. *Leuconostoc mesenteriodes* P-60 was used for methionine assay[13] and a cholineless *Neurospora*† mutant[14] for the choline assay. Assays were made on material before and after hydrolysis to preclude failure of detection of bound methionine. It was found that the particular preparation of Factor II used in these experiments contained 2γ of methionine per ml. The amount used brought in 0.8γ per ml. of medium. The L.E.L. brought in 0.4γ per ml. of medium. These amounts were included in calculations when methionine additions were made. In no case was the amount of methionine present lower than 0.8γ per ml. due to the necessity of adding the growth factor and where L.E.L. was also included there was never less than 1.2γ per ml. of medium.

Results.—It was found that homocystine could not replace methionine for *Tetrahymena* when the Factor II preparation was obtained from plant material. No growth resulted when methionine was withheld. This result was different from the one previously reported[6] when Factor II was prepared from Liver fraction L. The addition of varying amounts of L.E.L., however, resulted in growth where methionine was replaced by homocystine in the medium. The minimal concentration of the L.E.L. for moderately good growth (under our conditions) was found to be 120γ per ml. of medium and this amount was used throughout.

Dose response curves were constructed for homocystine. A typical curve is shown in figure 1 (curve 4). Increased growth was obtained until 40γ per ml. of homocystine were present. Higher amounts began to show inhibition. As a result of these experiments 40γ of DL-homocystine per ml. of medium were used routinely.

In view of the fact that 120 γ per ml. of L.E.L. made it possible for the ciliates to utilize homocystine in place of methionine, it was important to determine what effect the L.E.L. would have on growth in the presence of methionine. Accordingly two types of experiments were performed, one to determine the response to methionine alone and the other the response to

methionine plus 120γ per ml. of L.E.L. The dose response curves are shown in figure 1 (curve 2 for methionine and curve 3 for methionine plus L.E.L.). It will be noted that the minimum methionine present in curve 2 is 0.8γ per ml. (from the Factor II) while the minimum for curve 3 is 1.2γ per ml. (Factor II and L.E.L.). These curves show that the addition of L.E.L. with methionine makes no significant difference.

Cystine must be synthesized, under these conditions, from methionine and the addition of cystine had earlier been shown to spare methionine.[6] If this synthesis takes place via homocysteine formation as suggested by du Vigneaud[15] then the addition of homocystine to the medium should likewise spare methionine. This was found to be the case as shown in figure 1 (curve 1). The addition of 40γ of DL-homocystine per ml. spared approximately one half of the methionine. This addition made the 1.2γ of methionine, inevitably present as a minimum when L.E.L. was used, capable of supporting considerable growth (approximately 0.100 on the O.D. scale). It should be noted, however, that 0.8γ of methionine was insufficient for growth even in the presence of homocystine. It appears, therefore, that some of the action of the addition of L.E.L. to the homocystine medium is merely the bringing in of enough methionine to raise the level above the growth threshold. This does not fully account for the resulting growth when L.E.L. and homocystine replace methionine. Under these conditions, with only 1.2γ of methionine present, growth is more than double that obtained when L.E.L. is omitted (0.210 on O.D. scale). This difference must be accounted for by the conversion of homocystine to methionine.

Discussion.—The results outlined above confirm the observations of Toennies and Bennett [4,5] on the rat and point to an active agent in the liver fraction which in some way makes the conversion of homocystine to methionine possible. Labile methyl groups in the L.E.L. appear to be precluded in their experiments and in ours as the levels of methionine and choline could not account for enough to transmethylate the amount of homocystine used‡ even if all three methyl groups of choline were available. If the statement of McKittrick[3] regarding the number of methyl groups available from choline is considered, then the amount present is even less nearly adequate. The rat, unlike the chick, [3,16] can synthesize choline[17] provided methionine is present and we have found rapid synthesis of choline by *Tetrahymena* under like conditions.

Our results differ from those of du Vigneaud, *et al.* [1,2] and Toennies, *et al.*, [3,4] regarding choline plus homocystine. Unlike those workers we found that the addition of choline made no difference in the utilization of homocystine. Choline plus homocystine did not permit growth of the ciliates in the absence of adequate methionine or the L.E.L. agent. This fact indicates that the L.E.L. is not supplying a methyl donor but is supplying some agent concerned in the transmethylation process.

As a working hypothesis we make the following proposals to account for the action of the agent in L.E.L. in transmethylation; (I) In *Tetrahymena* the usual source of methyl for the synthesis of choline is methionine and this reaction is made possible by what we will call "transmethylase 1" (TM1). This enzyme appears to require no exogenous co-factor. The reaction may be diagrammed as follows, occurring in two steps:

$$\text{Serine} \longrightarrow \text{aminoethanol}^{18}$$

$$\text{Aminoethanol} + \text{methionine} \xrightarrow{TM1} \text{choline} + \text{homocysteine}.$$

(II) When methionine is replaced by homocystine, methionine is synthesized by the methylation of the latter, which is the reverse of the above reaction. In many cases one and the same enzyme is known to catalyze both a reaction and its reverse. The facts here do not support such activity on the part of TM1. This transmethylation (homocysteine to methionine) requires transmethylase 2 (TM2) which must be activated by an exogenous co-factor (CTM2) which the animal organism is incapable of synthesizing and therefore must obtain from its environment. This co-factor appears to be stored in the liver of mammals, presumably concentrated there from the food or from the products of microsynthesis by intestinal bacteria. It appears to be absent or in low concentrations in higher plants. Therefore, only when CTM2 (from L.E.L., in this case) is added with homocystine, is TM2 activated for the required transmethylation to methionine. (III) To account for the source of the methyl groups for this reaction one must assume the synthesis or mobilization of methyl from some unknown precursor (CH_3R), probably first to form choline and then to form methionine via homocysteine. (IV) The transmethylation of aminoethanol by CH_3R to form choline is presumably controlled by a transmethylating enzyme also and there is no evidence to indicate whether it is TM2 or a different one (TM3). It cannot be TM1 because no growth of the ciliate results in the presence of adequate choline and homocystine but no L.E.L. If it is a different enzyme from TM2 then it requires a different co-factor (CTM3) and this co-factor is also found in liver. This last assumption will become clear from the discussion of the results on the rat. The above reactions may be diagrammed as follows:

$$\text{Step 1.}\quad CH_3R \xrightarrow{\substack{TM2 + CTM2 \text{ (from L.E.L.)},\\ \text{or}\\ TM3 + CTM3 \text{ (from L.E.L.)},}} \text{choline}.$$

$$\text{Step 2.}\quad \text{Choline} + \text{homocysteine} \xrightarrow{TM2 + CTM2 \text{ (from L.E.L.)},} \text{methionine}$$

$$\text{Homocysteine} \longrightarrow \text{cysteine (sparing of methionine)}.$$

In the rat [1–5] step 2 seems to proceed without any exogenous co-factor but this can be explained on the basis of bacterial synthesis, storage or both. Both of these sources are denied *Tetrahymena* so step 2 only proceeds when CTM2 is added. In the rat step 1 is blocked by either too low a concentration of CTM2 for both steps or, if a different co-factor is required (CTM3), then its amount is inadequate within the body. In any case *Tetrahymena*, needing the co-factor for step 2, has a supply provided for step 1 by the addition of L.E.L.

The possibility of the direct methylation of homocysteine from CH_3R to form methionine should perhaps be considered. We know from the work of Simmonds, *et al.*,[19] that in the presence of added choline labeled methyl from choline does appear in the methionine. This need not be the only route of methionine formation, especially when choline, as a vitamin, is being synthesized. This alternative may be diagrammed as follows:

$$CH_3R + \text{homocysteine} \xrightarrow{TM2 + CTM2 \text{ (from bacterial synthesis + storage or from L.E.L.)}} \text{methionine}$$

$$\text{Methionine} + \text{aminoethanol} \xrightarrow{TM1} \text{choline (for vitamin activity).}$$

Recently, Horowitz[20] has presented evidence to show that the synthesis of methionine by *Neurospora* proceeds from cysteine and each step is controlled by one gene. One mutant failed to grow on homocysteine but did grow on methionine. It would be interesting and important to determine whether this failure to transmethylate is due to a failure on the part of the mutant to synthesize the necessary co-transmethylase. This could be determined by the addition of the active agent from L.E.L. The fact that none of the *Neurospora* mutants could utilize homocystine and all but the one mentioned above could grow on homocysteine is apart from the transmethylation problem, and concerns only the inability of the mold to reduce homocystine.[6]

The various parts of the above hypothesis are subject to test. Plans are made for the concentration and possible identification of the factor or factors in the liver extract. Factor II preparations entirely free of methionine and choline can be prepared so that the possible effects of traces of these substances can be determined. As concentration of the active agent or agents proceeds tests can be made to determine whether or not more than one co-factor exists which will indicate the relationships between the enzymes controlling steps 1 and 2.

Summary.—1. Homocystine can replace methionine in the diet of *Tetrahymena* provided some factor or factors found in liver are present.

2. The liver factor appears not to be a methyl donor as is indicated by

the failure of methionine replacement by homocystine in the presence of choline.

3. The liver factor can best be described as a co-transmethylase.

4. A discussion of the mechanism or mechanisms of transmethylation in animals is given.

* Aided by a grant from the U. S. Public Health Service and a grant recommended by the Committee on Growth acting for the American Cancer Society.

† The cholineless *Neurospora* used in this investigation was obtained from Dr. E. L. Tatum of Yale University.

‡ Betaine was absent from L.E.L. according to Bennett and Toennies.[5]

§ Since this manuscript went to press, we have tested strain 38706 (the methionineless mutant of *Neurospora crassa*, referred to as Me-1 by Horowitz[20]). Several other methionineless strains of *Neurospora* from the Yale collection (6073A, 4195a, 6279a, 4195A(3)) were also tested. All were found to be incapable of utilizing homocysteine, even in the presence of L.E.L. All responded to methionine. This indicates that the genetic block occurs in the synthesis of the transmethylating enzyme and not in the synthesis of the cofactor. We wish to thank Dr. N. H. Horowitz for strain 38706 and Drs. E. L. Tatum and S. Simmonds for the other strains and for the sample of homocysteine used.

[1] du Vigneaud, V., Chandler, J. P., Moyer, A. W., and Keppel, D. M., *J. Biol. Chem.*, **131**, 57 (1939).

[2] Moyer, A. W., and du Vigneaud, V., *Ibid.*, **143**, 373 (1942).

[3] McKittrick, D. S., *Arch. Biochem.*, **15**, 133 (1947).

[4] Toennies, G., Bennett, M. A., and Medes, G., *Growth*, **7**, 251 (1943).

[5] Bennett, M. A., and Toennies, G., *J. Biol. Chem.*, **163**, 235 (1946).

[6] Kidder, G. W., *Ann. N. Y. Acad. Sci.*, **49**, 99 (1947).

[7] Kidder, G. W., and Dewey, V. C., *Arch. Biochem.*, **6**, 425 (1945).

[8] Kidder, G. W., and Dewey, V. C., *Physiol. Zool.*, **18**, 136 (1945).

[9] Kidder, G. W., and Dewey, V. C., *Arch. Biochem.*, **8**, 293 (1945).

[10] Kidder, G. W., *Ibid.*, **9**, 51 (1946).

[11] Kidder, G. W., and Dewey, V. C., *Proc. Nat. Acad. Sci.*, **33**, 97 (1947).

[12] Kidder, G. W., and Dewey, V. C., *Ibid.*, **33**, 347 (1947).

[13] Dunn, M. S., Shankman, S., Camien, M. N., and Block, H., *J. Biol. Chem.*, **168**, 1 (1947).

[14] Horowitz, N. H., and Beadle, G. W., *Ibid.*, **150**, 325 (1943).

[15] du Vigneaud, V., Brown, G. B., and Chandler, J. P., *Ibid.*, **143**, 59 (1942).

[16] Jukes, T. H., *J. Nutrition*, **22**, 315 (1941).

[17] McKibbin, J. M., Thayer, B. A., and Stare, F. J., *J. Lab. Clin. Med.*, **29**, 1109 (1944).

[18] Stetten, D., Jr., *J. Biol. Chem.*, **144**, 501 (1942).

[19] Simmonds, S., Cohn, M., Chandler, J. P., and du Vigneaud, V., *Ibid.*, **149**, 519 (1943).

[20] Horowitz, N. H., *Ibid.*, **171**, 255 (1947).

LINKAGE AMONG GENES CONTROLLING INHIBITION OF LYSIS IN A BACTERIAL VIRUS*

BY A. D. HERSHEY AND RAQUEL ROTMAN†

DEPARTMENT OF BACTERIOLOGY AND IMMUNOLOGY, WASHINGTON UNIVERSITY SCHOOL OF MEDICINE, ST. LOUIS, MISSOURI

Communicated by M. Demerec, December 5, 1947

Experiments in which bacterial cells are simultaneously infected with two different viruses of the T2 serological group have suggested that both can grow in the same cell,[1] and that interaction between them produces new genetic types of virus.[2, 3] Further experiments, in which the mixed clones of virus coming from single bacterial cells were examined, have confirmed these inferences, and indicate that the new types arise only in those bacteria in which both viruses succeed in growing freely. These experiments are being continued in an effort to learn something about the mechanism of genetic interaction, and will not be reported in further detail at this time. In this paper we wish to describe a new series of mutations in the bacterial virus T2H,[1] discovered in connection with the experiments mentioned above, which point to the operation in this virus of a system of genetic linkage.

The virus T2H, like all its relatives, exists in the form of a lysis-inhibiting wild type (r^+) and an r (for rapidly lysing) mutant phenotype, which has lost the lysis-inhibiting character.[1] The principal new finding to be reported here is that all the independently arising r mutants of T2H that we have examined are genetically different. The basic observations may be summarized as follows.

A genetic difference between two r mutants is shown when bacteria infected with both liberate viral progeny containing a measurable proportion of lysis-inhibiting types, whereas bacteria infected with the same number of particles of either mutant alone yield only r viral progeny. The experimental procedure used is the one-step growth experiment of Delbrück and Luria, with mixed multiple infection, carried out as described earlier.[1] For simplicity of language, we shall speak of this experiment as a cross, described as heterologous when new types appear among the yield of virus, and homologous when only the parent types are recovered.

Certain details of the method require mention. First of all, one examines by it not the progeny of a genetic cross strictly speaking, but simply a mixed population of virus growing in the same community. The crosses reported in this paper were so arranged that the yield of virus examined came from about 40,000 mixedly infected bacteria. Variations in yields from individual bacteria[4] are thus left out of account. Care was taken to ensure adequate and equal multiplicity of infection, with respect both to titers of

stocks used and to adsorption to bacteria. Fortunately, all mutants were found to be equally well adsorbed in sensitive tests with admixed wild type (table 2). The effect of unequal multiplicity on the yield of segregants was also tested. The largest proportion of wild type resulted from equal multiplicity of infection with the two parent types, but the yield was not markedly affected by variations between 1 : 2 and 2 : 1. Variations in total multiplicity above about 3 of each type per bacterium had no measurable effect on the result. A multiplicity of about 5 of each type per bacterium was habitually used. Attention was also paid to the total yield of virus in crosses. In parallel experiments, the yields seldom differed by more than 15 per cent, and fell between 300 and 600 per bacterium, depending in part on variations in the nutrient broth. Because of these variations, difficult to control, important comparative results were always checked by parallel experiments.

The *r* mutants so far examined, among which 15 have been named *r*1, *r*2, etc., in order of isolation, are all identical with respect to type of plaque. Their behavior in intercrosses places them in two clearly defined classes, called A and B. Crosses between any pair belonging to the two classes yield about 15 per cent of wild type among the total yield of virus, no significant differences having been established. Crosses within classes yield proportions of wild type characteristic for each pair, falling between 0.5 and 8 per cent. The first eleven *r* mutants to be examined were intercrossed in all combinations, yielding only one example (*r*1) of class A, and no homologous pairs. Six others were examined less completely in a search for a second member of class A; all proved to be class B, and at least two differed from all the others. The second member of class A was finally obtained from an *r* mutant of the distinct virus T4, by crossing with T2H.[2] This mutant is called *r*14. Three examples of class B were discarded without numbering.

The results of the quantitative intercrosses, summarized in table 1, present three significant features. (1) The amounts of wild type appearing in the three possible crosses among sets of three mutants of class B (e.g., *r*2, *r*5 and *r*7) are approximately additive, but generally fall noticeably short, so that no linear order valid for several loci can be discerned. (2) The results of intercrosses among sets of four mutants (e.g., *r*2, *r*4, *r*5 and *r*7) cannot be interpreted in terms of four independent transfer frequencies, but indicate some type of linkage. (3) If linkage is assumed, the classes A and B are probably to be regarded as independent linkage groups. However, the fact that *r*1 of class A gives equal amounts of wild type with all members of class B is not adequate proof of this independence, for the same thing is seen at a lower level in the crosses between *r*13 and the other members of class B.

The different *r* mutants also differ with respect to rate of spontaneous

TABLE 1

PER CENT OF WILD TYPE IN VIRUS YIELDS FROM BACTERIA INFECTED WITH PAIRS OF *r* MUTANTS

CLASS A		CLASS B								
*r*1	*r*14	*r*13	*r*2	*r*4	*r*8	*r*9	*r*3	*r*5	*r*7	*r*6
*r*1	0.5	14	17	16	(15)	(13)	14	15	15	15
	*r*14	(13)	X	X	X	X	X	X	15	X
		*r*13	6	6	(7)	X	(7)	(6)	7	8
			*r*2	1.1	1.7	1.5	2.0	3.9	3.0	2.4
				*r*4	1.1	(1.5)	2.6	2.9	3.5	3.2
					*r*8	(0.8)	(2.1)	(2.3)	(2.3)	2.7
						*r*9	(0.9)	(1.8)	(2.2)	(2.4)
							*r*3	(1.2)	(1.0)	1.4
								*r*5	0.8	(1.1)
									*r*7	(0.5)

The crosses indicated by X have not been made. Results in parentheses are from single experiments; the others are averages of several. Self-crosses yield less than 0.1% wild type, usually none.

reversion to lysis-inhibitor during serial passage of the virus.[1] In two or more independent tests starting from different single plaques of each of several *r* mutants, the number of passages required to yield lysates containing 10 to 20 per cent of lysis-inhibiting virus varied as follows: for *r*9, two or three; for *r*5, *r*6, *r*11, five; for *r*2, *r*3, *r*4, *r*8, *r*10, *r*14, six or seven; for *r*7, *r*13, eight; for *r*1, ten; and for *r*15, about twenty-two. The lysis-inhibitor recovered from *r*2, *r*13, and *r*15 differed from authentic wild type in appearance of plaques, being of intermediate character between *r* and wild type. On further serial passage, variants apppeared which resembled wild type more closely. Both types of lysis-inhibiting virus obtained from these three mutants were genetically different from wild type, for when either was crossed with wild type, excessive numbers of *r* mutants were found among the progeny. This was true of seven independently arising reversions from one or another of the mutants named. Presumably the lysis inhibiting virus selected in these cases arose by suppressor mutations, rather than by back mutation at the respective *r* loci. On the other hand, lysis-inhibiting viruses recovered from each of the 11 remaining *r* mutants mentioned above were indistinguishable from wild type both phenotypically and in back cross. We have not, however, done the careful experiments which might reveal suppressor mutations closely linked to an *r* locus.

Two different *r* loci were found to mutate independently in a double-mutant virus obtained as described in the following section of this paper. In four independent serial passages starting from different single plaques of the double mutant *r*1*r*5, the reversion passed through the more stable type *r*1 in each case. This suggests that the difference in the rate of reversion of the two single mutants is due to differences in stability of loci, rather than to

different selection rates among mutants growing in competition with wild type. Direct estimations of selection rates have not been made, however.

In contrast with the spontaneous reversions, lysis-inhibiting virus arising in crosses between *r* mutants is always homologous with wild type. Tests of this point included segregants from the cross *r*2 × *r*13, two mutants whose spontaneous reversions are heterologous with wild type. There is, furthermore, no correlation between the individual stability of two *r* mutants and the amount of wild type they throw in crosses.

A few of the mutants, including two lines each of *r*2 and *r*13 mentioned above, were carried through the transformations $r \rightarrow hr \rightarrow r$ by crossing with a host-range mutant (*h*), followed by a cross with wild type. The resulting stocks showed the same behavior in crosses as the originals. It is unlikely, therefore, that these are multiple-locus mutants.

In order to compare the ability of the different *r* mutants to grow competitively in the bacterial cells, mixed infections with *r* mutant and wild type were made. In only two cases did the proportion of wild type in the

TABLE 2

COMPETITION BETWEEN *r* MUTANTS AND WILD TYPE IN MIXED INFECTION

MUTANT	% WILD TYPE			MUTANT	% WILD TYPE		
	INPUT	ADSORBED	YIELD		INPUT	ADSORBED	YIELD
*r*1	54	59	53	*r*4*r*7	52	55	53
*r*5	50	57	54	*r*2	68	77	64
*r*1*r*5	53	55	37	*r*3	58	58	60
*r*13	51	48	60	*r*6	60	60	57
*r*4	60	64	65	*r*2*r*3	63	70	67
*r*7	59	54	61	*r*2*r*6	60	56	64
*r*4*r*7*r*13	59	54	54	*r*3*r*6	66	71	69

yield differ appreciably from the proportion in the infecting mixture, or from the proportion in the fraction adsorbed to bacteria. With the mutant *r*13 the yield contained a proportion of wild type in excess of that adsorbed amounting to about 10 per cent of the total virus (3 experiments). The significance of this is not clear. The other exception is easily explained. The cross between the double mutant *r*1*r*5 and wild type yielded an increased proportion of *r* mutant amounting to about 15 per cent of the total virus (2 experiments), owing to the segregation of the two *r* loci described below. For other multiple mutants studied, all of class B, the loss of wild type was too small to be detected, presumably because of linkage between the loci. The data for individual experiments are summarized in table 2.

Segregation of Loci.—The cross between any two *r* mutants yields two new viral types: wild type and an *r* mutant distinguishable from others only by back cross to the parental types, in which it proves homologous with both. This behavior supports the idea that each mutant represents a modification at a different genetic locus among many controlling inhibition of lysis.

For any pair X and Y, the segregation may be written: $(X^+) \times (^+Y) \rightleftharpoons (^{++}) \times (XY)$. In testing this hypothesis, it is necessary to prepare single-plaque stocks from a number of the r progeny of the cross, and test each one by back cross to each of the parent types. Tests of this kind have confirmed the hypothesis in every case. It was also found, within the limits imposed by the small scale of the experiments, that in each cross the two recombination types appeared in equal numbers, and the two parental types in the same proportion with which the bacteria were initially infected. These data are shown in table 3. In one experiment, virus yields from single bacteria[4] infected with *r*4 and *r*7 were found to contain a small proportion of wild type in nearly every burst, the average being about 4 per cent. From one of the yields, tests of a number of *r* progeny revealed the three additional viral types expected.

TABLE 3

SEGREGATION OF *r* LOCI IN INTERCROSSES

CROSS	WILD TYPE (% OF YIELD)	DISTRIBUTION OF TYPES AMONG *r* PROGENY			
*r*1 × *r*5	15	4 *r*1	4 *r*5	2 *r*1*r*5	
*r*1 × *r*7	15	9 *r*1	5 *r*7	6 *r*1*r*7	
*r*1 × *r*13	15	5 *r*1	3 *r*13	4 *r*1*r*13	
*r*4 × *r*13	6	12 *r*4	17 *r*13	1 *r*4*r*13	
*r*7 × *r*13	7	7 *r*7	7 *r*13	1 *r*7*r*13	
*r*2 × *r*3	3.3	36 *r*2	47 *r*3	3 *r*2*r*3	
*r*2 × *r*6	3.4	33 *r*2	41 *r*6	1 *r*2*r*6	
*r*3 × *r*6	1.9	63 *r*3	23 *r*6	1 *r*3*r*6	
++ × *r*1*r*5	37	2 *r*1	2 *r*5	6 *r*1*r*5	
*r*1*r*5 × *r*1*r*13	0.0	13 *r*1*r*5	13 *r*1*r*13	2 *r*1*r*5*r*13	2 *r*1
*r*4*r*7 × *r*7*r*13	0.0	12 *r*4*r*7	15 *r*7*r*13	3 *r*4*r*7*r*13	0 *r*7

The equality of numbers of the two recombination progeny, only suggested by the data of table 3, has been confirmed on a statistically adequate scale for crosses between host-range and *r* mutants to be reported elsewhere.

Multiple Factor Tests.—In an effort to learn something about the nature of the linkage, multiple factor crosses have been made. These are of three types.

The first type consists of three factor crosses involving the loci *r*1, *r*4, and *r*7, the two latter being rather closely linked but probably independent of the first. If this is correct, one would expect the factor *r*1 to distribute itself equally between the two types of virus resulting from the enchange of *r*4 with respect to *r*7. The effect should be an equal segregation of wild type from the crosses *r*1*r*4 × *r*7 and *r*1*r*7 × *r*4, amounting in either case to about half that segregating in the cross *r*4 × *r*7. For the cross *r*1 × *r*4*r*7, the yield of wild type should be nearly as large as that for *r*1 × *r*7, owing to the linkage. The data shown in table 4 confirm these expectations very well.

TABLE 4

SEGREGATION OF WILD TYPE IN MULTIPLE FACTOR INTERCROSSES

CROSS	% WILD TYPE	CROSS	% WILD TYPE
r1r4 × *r7*	1.4 (3)	*r7r13* × *r4r13* × *r4r7*	0.16 (2)
r1r7 × *r4*	1.4 (3)	*r4* × *r7* × *r13*	7.3 (1)
r1 × *r4r7*	11 (2)	*r2r3* × *r2r6* × *r3r6*	0.02 (7)
r1 × *r7*	15 (4)	*r2* × *r3* × *r6*	3.0 (2)
r4r13 × *r7*	1.8 (4)	*r2r3* × *r6*	0.8 (7)
r7r13 × *r4*	0.7 (3)	*r2r6* × *r3*	0.3 (6)
r4r7 × *r13*	5.0 (3)	*r3r6* × *r2*	1.4 (5)
r4 × *r13*	6.3 (5)	*r2* × *r3*	2.0 (6)
r7 × *r13*	7.0 (8)	*r2* × *r6*	2.4 (6)
r4 × *r7*	3.5 (9)	*r3* × *r6*	1.4 (7)

The results shown are averages of the number of experiments indicated in parentheses. The 22 measurements, including all crosses *r4* × *r13*, *r7* × *r13* and *r4* × *r7* made during several months, show a coefficient of variation of 20%.

The second type of cross corresponds to the conventional three factor test for linkage order. We have used two sets of three factors all belonging to the class B. The sets include (*r2*, *r3* and *r6*) and (*r13*, *r4* and *r7*), the former being more closely linked than the latter. The data for these crosses, also included in table 4, tend to place the factors in the order listed above, though the correspondence with expectations for random crossing over between linear structures is rather poor. The data for *r2*, *r3* and *r6*, particularly, are compatible with the possibility that a crossover in one region tends to be accompanied by a second in the adjacent region.

The third type of cross, suggested by analogous experiments with bacteria,[5] was tried for the purpose of estimating the amount of repeated exchange (i.e., exchange between recombination progeny and progeny of the parental types) taking place under the conditions of the experiments. The combinations used were such that segregation of wild type would require the interaction (simultaneous or successive) of three different mutants. The sets chosen were (*r4r13*, *r7r13*,*r4r7*) and (*r2r3*, *r2r6*, *r3r6*). In different experiments, multiplicities of about three and five per bacterium of each of the three mutants of the set were tested, with similar results. The data for these crosses, and for the corresponding sets of three single mutants, are shown in table 4. It will be seen that repeated exchanges can be detected by this method; that they are less frequent for the more closely linked factors; and that their number is too small to be of quantitative importance in the linkage tests already described.

It should be noted that repeated exchange of the type detected above should not have any effect in two factor crosses. Repeated exchange between sister recombination progeny, if such arise, presents a more serious difficulty. Until effects from this cause can be evaluated, no great quantitative significance can be attached to the results of intercrosses.

Discussion.—Differences among the *r* mutants described in this paper indicate that there are an indefinite number of distinct loci controlling lysis inhibition in the bacterial virus T2H. The suggestion from the genetic data is that the mechanism of lysis-inhibition[6] may be exceedingly complex. On the other hand, the fact that identical mutants are not encountered shows that the different *r* mutations occur with similar frequency. This remarkable circumstance might be interpreted as an indication that the different mutations result from a single type of chemical change occurring at different points in the genetic material of the viral particle. If this were so, the different *r* mutants might represent alterations of the same biochemical property. This argument must be qualified in view of our ignorance of the number of possible *r* mutations.

The existence of suppressor mutations noted in this paper, and of intermediate mutant types,[3] complicates the system somewhat further.

It is evident that the rather high frequency of *r* mutations, about 10^{-4} per viral duplication,[1] is at least partly due to the large number of loci involved. It is probably not feasible to measure the stability of any given *r* locus. For this reason the test of equal rates of *r* mutation in wild type and in the identical phenotype obtained by reversion from *r*,[3] is of no value as a test of reversibility of mutation. The test by crossing with wild type, reported in this paper, shows that some reversions are not back mutations, but that most of them probably are.

The situation described for the *r* mutants does not prevail among viral mutants in general. We have examined six independently arising host-range mutants of the h^c phenotype,[3] which proved to be homologous in intercrosses. This locus can be assigned a position in the linkage class B however, and it is for their usefulness in mapping that the *r* mutants are likely to prove most valuable.

The linkage system itself we shall discuss only briefly, pending the completion of experiments of another type. It is clear that the clases A and B are independent, or nearly so, and that some kind of linkage exists within the classes. Regarding the linkage structure, some of our data suggest linear arrangement with crossing over, but others are somewhat contradictory. Among difficulties which cannot be discussed very intelligently at present, the necessity for reconciling our data with the results of Luria[7] should be mentioned. It is probably necessary to add that the linkage system discerned in the virus may yet prove to have little in common with the more familiar linkage systems in cellular organisms.

Summary.—Different *r* mutants of the virus T2H differ (*a*) in rate of spontaneous reversion to lysis-inhibiting types; (*b*) in the kind of reversion they undergo, either to a form genetically like, or to forms unlike, the wild type; (*c*) in the amount of wild type they throw in intercrosses. With respect to (*c*) there are two independent linkage groups of distinct *r* loci.

Analysis of one of these groups by single and multiple factor intercrosses has failed to establish or refute the idea of crossing over between linear structures. The existence of some kind of linkage system conditioning segregation and reassortment of genetic factors among viral particles seems quite clear, but it remains to be learned whether reciprocal, or even material, exchanges are involved.

* Aided by a grant from the United States Public Health Service.

† On leave from the Instituto Bacteriologico de Chile, Santiago, Chile.

[1] Hershey, A. D., *Genetics*, **31**, 620–640 (1946).

[2] Delbrück, M., and Bailey, W. T., Jr., *Cold Spring Harbor Symp. Quant. Biol.*, **11**, 33–37 (1946).

[3] Hershey, A. D., op. cit., 67–76.

[4] Delbrück, M., *J. Bact.*, **50**, 131–135 (1945).

[5] Lederberg, J., "Gene Recombination and Linked Segregations in *Escherichia coli*," *Genetics*, **32**, 505–525 (1947).

[6] Doermann, A. H., see "Discussion" in reference 3, and complete paper to appear in *J. Bact.* **55**, 257–276 (1948).

[7] Luria, S. E., these PROCEEDINGS, **33**, 253–264 (1947).

---

*A THEORY OF MINIMAX**

BY MAX SHIFFMAN

NEW YORK UNIVERSITY

Communicated by S. Lefschetz, January 7, 1948

1. Variational analysis, or the theory of critical points, can be examined either analytically or topologically. To furnish the link between these two aspects, it is necessary to establish certain basic properties of the functional in question beyond what is called for in the case of the pure minimum. Lower semi-continuity and compactness are adequate for minima. But precisely this semi-continuity, and not continuity, introduces new questions in the more general case of minimax. A link between the analytic and the topologic is established if the functional can be shown to be *reducible*. This link is completed if it can be shown that the minimax type defined topologically is equal to the minimax type defined analytically. Only when these facts are established is it possible to pass between the domains of analysis and of topology.

In the present note, we shall sketch a theory which does not separate these two basic questions, reducibility and type, but rather makes an organic unity of them (Theorem 3). We discuss a single integral problem where the functional ranges over curves, but it will be clear that the theory is in principle applicable to problems involving surfaces or higher dimensional varieties. Indeed these latter problems are the main motive and

goal of the present work and will be considered in the detailed publication.

2. Consider the variational problem for

$$I[y] = \int_{x_0}^{x_1} f(x, y, y')dx \tag{1}$$

with prescribed end-points $y(x_0) = y_0$, $y(x_1) = y_1$. Let $y = \varphi(x)$ be an extremal passing through these points. We shall suppose that f is sufficiently often differentiable in its arguments x, y, y' and that

$$f_{y'y'}(x, y, p) > 0 \tag{2}$$

for all y in a neighborhood of $\varphi(x)$ and for all p, i.e., we suppose that (1) is a regular variational problem near $y = \varphi(x)$. The second variation of $I[y]$ is

$$\delta^2 I[\eta] = \int_{x_0}^{x_1} [f_{yy}\eta^2 + 2f_{yy'}\eta\eta' + f_{y'y'}\eta'^2]dx \tag{3}$$

where the arguments in the terms involving f are $(x, \varphi(x), \varphi'(x))$ and where

$$\eta(x_0) = \eta(x_1) = 0. \tag{4}$$

The quadratic functional (3) subject to any normalizing condition, say $\int \eta^2 dx = 1$, has certain successive eigenvalues and the corresponding eigenfunctions. Select the first n of them, $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $u_1(x), u_2(x), \ldots, u_n(x)$, respectively, where n is chosen so that

$$\lambda_{n+1} > 0. \tag{5}$$

The eigenfunctions $u_i(x)$ satisfy

$$(f_{yy}u_i + f_{yy'}u_1') - \frac{d}{dx}(f_{y'y}u_i + f_{y'y'}u_i') - \lambda_i u_i = 0$$

$$u_i(x_0) = u_i(x_1) = 0 \tag{6}$$

and are normalized so that

$$\int_{x_0}^{x_1} u_i(x)u_j(x)dx = \delta_{ij}. \tag{7}$$

With these preliminaries, the following fundamental theorems of the theory can be stated (for Theorems 1, 2 a much weaker hypothesis than (2) can be made).

THEOREM 1. *There is a family of functions*

$$y = \varphi(x; c_1, \ldots, c_n) \tag{8}$$

depending on the n parameters $c_1, \ldots, c_n$ and defined for sufficiently small absolute values of these parameters, which lie in some neighborhood in the (x, y) plane of $y = \varphi(x)$, which reduces to $y = \varphi(x)$ for $c_1 = \ldots = c_n = 0$, and which passes through the end-points (x_0, y_0), (x_1, y_1). For any curve

$y = y(x)$ passing through the end-points and lying in this neighborhood of $y = \varphi(x)$, determine the values of $c_1, \ldots, c_n$ by

$$\int_0^{x_1} [y(x) - \varphi(x)]u_i(x)dx = c_i, \qquad i = 1, \ldots, n \tag{9}$$

Then

$$I[y(x)] \geqq I[\varphi(x;\ c)]. \tag{10}$$

THEOREM 2. *For the family $\varphi(x;\ c)$ we have*

$$I[\varphi(x\ c)]; = I[\varphi(x)] + \sum_{i=1}^{n} \lambda_i c_i^2 + r \tag{11}$$

where the remainder term r is estimated by

$$\frac{|r|}{c_1^2 + \ldots + c_n^2} \to 0 \text{ as } c_1, \ldots, c_n \to 0$$

uniformly in $c_1, \ldots, c_n$.

THEOREM 3. *Make the additional assumption that there are positive numbers A, M such that*

$$\left|\frac{f_y(x, y, p)}{pf_{y'}(x, y, p)}\right| \leqq M \text{ for all } |p| \geqq A \tag{12}$$

and for all (x, y) in a neighborhood of $y = \varphi(x)$. Consider the quantity

$$I(t) = I[\varphi(x;\ c) + t\{y(x) - \varphi(x;\ c)\}] \tag{13}$$

as a function of t in $0 \leqq t \leqq 1$, where $c_1, \ldots, c_n$ are determined from (9). Then

$$\frac{dI(t)}{dt} \geqq \frac{1}{2}\frac{I(t) - I(0)}{t} \geqq 0$$

and

$$I(t) \leqq I(0) + t^{1/2}\{I(1) - I(0)\}.$$

3. To prove Theorems 1, 2, consider the problem of minimizing $I[y(x)]$ among curves passing through (x_0, y_0), (x_1, y_1) and satisfying the side conditions (9) for fixed but small values of $c_1, \ldots, c_n$. The extremals of this new variational problem are solutions of

$$f_y - \frac{d}{dx} f_{y'} - \sum_{i=1}^{n} \mu_i u_i(x) = 0 \tag{14}$$

where $\mu_1, \ldots, \mu_n$ are constant Lagrange multipliers, and they must also satisfy (9). Let

$$Y(x;\ b, \mu_1, \ldots, \mu_n) \tag{15}$$

be solutions of (14) with initial conditions

$$Y(x_0, b, \mu) = y_0$$

$$Y_x(x_0, b, \mu) = b + \sum_{i=1}^{n} \frac{\mu_i}{\lambda_i} u_i'(x_0). \tag{16}$$

A particular solution of (14), (16) with $\mu_1 = \ldots = \mu_n = 0$ and $b = \varphi'(x_0)$ is $\varphi(x)$.

We establish first that

$$Y_{\mu_i}(x, \varphi'(x_0), 0) = \frac{1}{\lambda_i} u_i(x). \tag{17}$$

Then insert (15) for $y(x)$ in (9) and consider the resulting equations as n equations for the determination of $\mu_1, \ldots, \mu_n$ in terms of $b, c_1, \ldots, c_n$. For $b = \varphi'(x_0)$, $c_1 = \ldots = c_n = 0$, a solution of these equations is $\mu_1 = \ldots = \mu_n = 0$. By virtue of (17), we prove that the Jacobean of these equations, evaluated for $b = \varphi'(x_0)$, $c_1 = \ldots = c_n = 0$, is $\neq 0$. Therefore we can solve the equations for $\mu_1, \ldots, \mu_n$ in terms of $b, c_1, \ldots, c_n$ for b sufficiently near $\varphi'(x_0)$ and $c_1, \ldots, c_n$ sufficiently small in absolute value. Substituting in (15) gives a function

$$Z(x, b, c_1, \ldots, c_n) \tag{18}$$

which is an extremal and satisfies (9).

We establish now that the equation

$$Z(x_1, b, c) = y_1$$

has a unique solution for b in terms of $c_1, \ldots, c_n$ for sufficiently small values of these parameters. Substituting in (18) gives the final function

$$y = \varphi(x; c_1, \ldots, c_n) \tag{19}$$

which is an extremal, passes through the given points and satisfies (9).

For the extremal (19), the second variation of the new variational problem is

$$\int_{x_0}^{x_1} \{\bar{f}_{yy}\eta^2 + 2\bar{f}_{yy'}\eta\eta' + \bar{f}_{y'y'}\eta'^2\}dx \tag{20}$$

with

$$\int_{x_0}^{x_1} \eta u_i(x)dx = 0, \qquad i = 1, \ldots, n, \tag{21}$$

and (4). The bar over the quantities involving f in (20) indicate that the argument is $(x, \varphi(x; c), \varphi'(x; c))$. The quadratic functional (3) subject to (21) and (4) is positive definite, and has a value $\geqq \lambda_{n+1} \int \eta^2 dx$. Since $\varphi(x; c)$ and $\varphi'(x; c)$ are near $\varphi(x)$, $\varphi'(x)$, the second variation (20) subject to (21) and (4) is also positive definite. The positive-definiteness of the second variation implies that the extremal (19) can be imbedded in a

suitable field of extremals. Furthermore, by virtue of the condition (2) and the fact that the side conditions (9) do not involve y', the Weierstrass E-function for this variational problem is positive. Therefore (19) actually minimizes, and this proves Theorem 1.

Theorem 2 follows from the method of determination of $\varphi(x;\ c)$. For we prove that

$$\varphi_{c_j}(x;\ 0,\ \ldots,\ 0) = u_j(x),$$

and this will yield Theorem 2.

Incidentally in (16) we tacitly assumed that none of the λ_i are zero. If some $\lambda_i = 0$, a suitable modification of (16) can be made.

4. To prove Theorem 3, consider the integral

$$J[\omega] = \int_{x_0}^{x_1} g(x,\ \omega,\ \omega')dx \tag{22}$$

where

$$\begin{aligned} g(x,\ \omega,\ \omega') = {} & \omega f_y(x,\ \varphi(x;\ c) + \omega,\ \varphi'(x;\ c) + \omega') \\ & + \omega' f_{y'}(x,\ \varphi(x;\ c) + \omega,\ \varphi'(x;\ c) + \omega') \\ & - {}^1/_2 f(x,\ \varphi(x;\ c) + \omega,\ \varphi'(x;\ c) + \omega') \end{aligned} \tag{23}$$

and where $\omega(x)$ is subject to the side conditions

$$\int_{x_0}^{x_1} \omega u_i(x)dx = 0, \qquad i = 1,\ \ldots,\ n \tag{24}$$

and the end conditions

$$\omega(x_0) = \omega(x_1) = 0. \tag{25}$$

The variational problem for $J[\omega]$ subject to (24), (25), which we shall call problem J, has as Euler equation

$$g_\omega - \frac{d}{dx}\, g_{\omega'} - \sum_{i=1}^{n} \mu_i u_i(x) = 0. \tag{26}$$

But for the curve $\omega = 0$, we have

$$g_\omega = {}^1/_2 \bar{f}_y, \qquad g_{\omega'} = {}^1/_2 \bar{f}_{y'}$$

where the bar over the f indicates that the argument is $(x,\ \varphi(x;\ c),\ \varphi'(x;\ c))$. Since $\varphi(x;\ c)$ satisfies (14) for suitable values of $\mu_1,\ \ldots,\ \mu_n$, it follows that $\omega(x) \equiv 0$ satisfies (26) for these constants $\mu_1,\ \ldots,\ \mu_n$ and so

$$\omega(x) \equiv 0 \tag{27}$$

is an extremal for problem J.

Evaluating for the extremal (27), we have

$$g_{\omega'\omega'} = {}^3/_2 \bar{f}_{y'y'} > 0. \tag{28}$$

The second variation of J is likewise ${}^3/_2$ times the expression (20), with

side conditions (21). The extremal can therefore be imbedded in a field of extremals for the problem J.

The Weierstrass E-function

$$E = g(x, \omega, \sigma) - g(x, \omega, \omega') - (\sigma - \omega')g_{\omega'}(x, \omega, \omega') \qquad (29)$$

for ω, ω' in the field, and σ arbitrary, can be shown to be positive in a sufficiently small neighborhood of (27). For from (28), there is a constant a such that $E > 0$ for $|\sigma| \leqq a$ (and $\sigma \neq \omega'$). By virtue of (2) and (12), there is a constant B such that $E > 0$ for $|\sigma| \geqq B$. Finally, in the range $a \leqq |\sigma| \leqq B$ we can make $E > 0$. In all three cases, sufficiently small neighborhoods of (27) in the field must be taken. This establishes the fact that (27) actually minimizes $J[\omega]$ in the final neighborhood selected. Hence

$$J[\omega(x)] \geqq -\tfrac{1}{2}I[\varphi(x;\ c)]. \qquad (30)$$

Consider the quantity $I(t)$ defined in (13). We have

$$t\frac{dI(t)}{dt} - \frac{1}{2}I(t) = J[\omega] \qquad (31)$$

where

$$\omega = t\{y(x) - \varphi(x;\ c)\}.$$

Inequality (30) yields

$$t\frac{dI(t)}{dt} - \frac{1}{2}I(t) \geqq -\frac{1}{2}I(0).$$

This is the first inequality of Theorem 3. The second inequality is an immediate consequence of the first.

5. Theorem 3 can be formulated as a fundamental deformation theorem which establishes the reducibility of $I[y]$ and the characterization of the minimax type of the extremal $y = \varphi(x)$. Select n so that λ_{n+1} is the first positive eigenvalue. There is a neighborhood of $y = \varphi(x)$ such that any curve $y(x)$ in this neighborhood can be deformed into $\varphi(x;\ c)$ in such a way that the functional $I[y]$ always decreases, and at a rate given in a uniform way in Theorem 3. The end family $\varphi(x;\ c)$ is an n-parametric family for which Theorem 2 holds. This proves reducibility and at the same time shows, in case none of the λ_i are zero, that the topologic type of minimax is exactly equal to the number of negative eigenvalues of the associated boundary value problem (6).

* Dedicated to Richard Courant on his 60th birthday.

ON ALL OF MERSENNE'S NUMBERS PARTICULARLY M_{193}

BY H. S. UHLER

YALE UNIVERSITY

Communicated by F. D. Murnaghan, January 15, 1948

On November 27, 1947, the author finished investigating, by application of the Lucasian sequence 4, 14, 194, 37634, $\cdots$, the factorizability of Mersenne's number $M_{193} = 2^{193} - 1 =$ 1255 42034 70773 36152 76715 78846 41533 28322 04710 88892 80690 25791. The 192nd residue had the value 542 45701 25193 90814 13211 43009 56802 04633 04970 79432 42801 51282 which, being non-zero, shows[1] that M_{193} is composite in character. A sufficiently detailed account of the procedure followed and of the underlying basic theorem may be found in the author's first paper[2] on Mersenne's numbers. Suffice it to record two facts. After the date given above the data of the entire set of original work-strips were checked for the third time with an auxiliary modulus in conformity with the condition $(M_{193}q_k + r_k + 2) \equiv r_{k-1}^2 \pmod{10^7 + 1}$, where q and r denote "quotient" and "remainder," respectively. The approximation to the reciprocal of M_{193} used in the latest work had the value 0. (58 zeros) 79654 59555 66226 13851 44401 98883 85590 27955 52277 59630 93930 37006 37523 90143 41987 11093 10854 40700 94876 19334 41734 25085 25958 34273 96290 7$\cdots$. The check multiplication gave $(M_{193})(1/M_{193})_a = 1 + 1.98 \times 10^{-126}$.

With regard to the 55 numbers of the form $2^p - 1$ (where p is a prime not exceeding 257) considered by Mersenne, the following brief report on recent progress made in their study and the present state of our knowledge may be timely because of its definitive character. In the year 1935 there remained[3] six Mersenne numbers which had not been investigated with respect to their prime or composite properties. These corresponded to $p = 157, 167, 193, 199, 227$ and 229.

In leisure hours the present writer began the investigation of M_{157} in the spring of the year 1944 and finished the entire set with M_{193} on the date given above. These M_p's were taken up in random order in the vain hope of discovering a prime number greater than $2^{127} - 1$. More specifically the final residues for M_{157}, M_{167}, M_{229}, M_{199} and M_{227} were obtained, respectively, on Aug. 11, 1944, Dec. 2, 1944, Feb. 9, 1946, July 27, 1946, and June 4, 1947. (The manuscript on M_{227} was accepted for publication in the *Bulletin of the American Mathematical Society* on July 17, 1947.) Since all six M_p's were found to be composite the earlier status of Mersenne's remarks has not been changed but a lacuna in our knowledge has been filled. The next extremely difficult step will consist in the discovery of all of the as yet unknown factors of Mersenne's numbers.

Mersenne said[3] that the only values of p not greater than 257 which

make M_p prime are 2, 3, 5, 7, 13, 17, 19, 31, 67, 127 and 257. Comparison of this list with the correct data recorded in the top line of the table presented below shows that Mersenne made five mistakes. $p = 67$ and 257 do not yield prime values for M_p, and $p = 61$, 89 and 107 were not included in his list of special primes.

With reference to explicit factoring, attention should be called to a valuable paper[4] by Professor D. H. Lehmer entitled "On the Factors of $2^n \pm 1$." His investigations on 76 numbers unveiled eleven factors which fall within Mersenne's range. Incidentally two of his new factors confirmed the present writer's final residues for M_{167} and M_{229}.

p	CHARACTER OF M_p
2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127	Prime
11, 23, 29, 37, 41, 43, 47, 53, 59, 67, 71, 73, 79, 113	Composite and fully factored
151, 163, 173, 179, 181, 223, 233, 239, 251	Two or more prime factors found
83, 97, 131, 167, 191, 197, 211, 229	Only one prime factor known
101, 103, 109, 137, 139, 149, 157, 193, 199, 227, 241, 257	Composite but no factor known

[1] Lehmer, D. H., *Jour. London Math. Soc.*, **9–10**, 162–165 (1934–1935).
[2] Uhler, H. S., *Proc. Nat. Acad. Sci.*, **30**, 314–316 (1944).
[3] Archibald, R. C., *Scripta Mathematica*, **3**, 112–119 (1935).
[4] Lehmer, D. H., *Bull. Amer. Math. Soc.*, **53**, 164–167 (1947).

NEW TYPES OF CONGRUENCES INVOLVING BERNOULLI NUMBERS AND FERMAT'S QUOTIENT

BY H. S. VANDIVER

DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF TEXAS

Communicated January 17, 1948

In another paper[1] the writer gave the relation

$$(mb + k)^n = \sum_{a=1}^{r} \binom{r}{a} (-1)^{a-1} \frac{S_n(m, k, a)}{a}; \qquad r > n; \qquad (1)$$

where we define the Bernoulli numbers b_n by means of the recursion formula $(b + 1)^n = b_n$; $n > 1$, the left-hand member being expanded in full and b_s substituted for b^s, and the left-hand member of (1) is interpreted in the same way;

$$S_n(m, k, a) = \sum_{i=0}^{a-1} (im + k)^n; \qquad 0^0 = 1,$$

m and k are any integers with $m \neq 0$. In the present paper we shall employ (1) as well as other known relations to obtain various congruences. The principal results proved seem to be (5), (6b), (7) and (20).

By the well-known Bernoulli summation formula we also have

$$(b + k)^n = nS_{n-1}(k) + b_n, \tag{2}$$

where

$$S_{n-1}(k) = S_{n-1}(1, 0, k).$$

Set in (1), $r = p - 1$ we have, if $n < p - 1$, p being an odd prime,

$$(mb + k)^n = \sum_{a=1}^{p-1} \binom{p-1}{a} (-1)^{a-1} \frac{S_n(m, k, a)}{a}. \tag{3}$$

It is known that

$$\binom{p-1}{a} \equiv (-1)^a \pmod{p}, \tag{3a}$$

so that from (3) we obtain

$$(mb + k)^n \equiv -\sum_{a=1}^{p-1} \frac{S_n(m, k, a)}{a} \pmod{p}. \tag{4}$$

For a particular $a = i$, the corresponding term on the right may be written

$$-\sum_{i=1}^{p-1} \frac{k^n + (k + m)^n + \ldots (k + (i - 1)m)^n}{i} = C.$$

Now collect the terms in the right-hand member of (4) involving each particular nth power, we find

$$C = -\sum_{s=0}^{p-2} (k + ms)^n \left(\frac{1}{s+1} + \frac{1}{s+2} + \ldots + \frac{1}{p-1}\right).$$

But also

$$1 + \frac{1}{2} + \ldots + \frac{1}{p-1} \equiv 1 + 2^{p-2} + \ldots + (p-1)^{p-2} \equiv 0 \pmod{p}, \tag{4a}$$

since p is odd. The last result gives

$$-\left(\frac{1}{s+1} + \frac{1}{s+2} + \ldots + \frac{1}{p-1}\right) \equiv 1 + \frac{1}{2} + \ldots + \frac{1}{s},$$

so that (4) may be written, $n < p - 1$,

$$(mb + k)^n \equiv \sum_{a=1}^{p-2} (ma + k)^n R_a \pmod{p}, \tag{5}$$

where

$$R_a = 1 + \frac{1}{2} + \ldots + \frac{1}{a}.$$

For $k = 0$, $m = 1$, this becomes

$$b_n \equiv \sum_{a=1}^{p-2} a^n R_a \pmod{p}. \tag{5a}$$

Also, (5) gives for g any integer, $0 < g < p$, for $mg + k \not\equiv 0 \pmod{p}$,

$$\begin{aligned}
&\sum_{n=0}^{p-2} (mg + k)^{-n+p-2} (mb + k)^n \\
&\equiv \sum_{a=1}^{p-1} R_a \sum_{n=0}^{p-2} (ma + k)^n (mg + k)^{p-2-n} \\
&\equiv -(mg + k)^{p-2} R_g + \sum_{\substack{a=1 \\ a \neq g}}^{p-1} R_a \frac{(ma + k)^{p-1} - (mg + k)^{p-1}}{m(a - g)},
\end{aligned}$$

or

$$\sum_{n=0}^{p-2} (mg + k)^{-n+p-2} (mb + k)^n \equiv -(mg + k)^{p-2} R_g, \tag{5b}$$

modulo p, by Fermat's theorem, provided $mg + k \not\equiv 0 \pmod{p}$. For $m = 1$, $k = 0$ this reduces to

$$\sum_{n=0}^{p-2} g^{-n} b_n \equiv -R_g \pmod{p}, \tag{6}$$

which for $g = 1$ becomes[2]

$$b_0 + b_1 + \ldots + b_{p-2} \equiv -1 \pmod{p}. \tag{6a}$$

Taking the relation (5a) for $n = 0, 1, \ldots, p - 2$ in turn introducing powers of the integers m, prime to p, similarly to the procedure in the proof of (5b) we obtain

$$\sum_{n=0}^{p-2} g^{-n} m^n b_n \equiv -R_a \pmod{p}, \tag{6b}$$

where $am \equiv g \pmod{p}$; $\quad a < p$.

Now take (5a) again and let $p = 1 + cm$ and put $n, n + c, \ldots, n + (m - 1)c$, in turn in this congruence, and add, we obtain if $n < c$

$$\sum_{i=0}^{m-1} b_{n+ic} \equiv m \sum_r r^n R_r \pmod{p}, \tag{7}$$

where r ranges over the distinct solutions in the set $1, 2, \ldots, p - 1$ of $x^c \equiv 1 \pmod{p}$. The relation (7) has a type of analogy with the relation[3]

$$\frac{\sum r^n}{p} \equiv -cn \sum_{k=0}^{m-1} \frac{b_{kc+n}}{kc + n} \pmod{p}, \tag{8}$$

with c and n even. Set $\mu_k \equiv k/m \pmod{p}$; $0 < \mu_k < p$. Then

$$(mb + k)^n = m^n \left(b + \frac{k}{m}\right)^n \equiv m^n(b + \mu_k)^n$$

$$\equiv m^n(b_n + n(0^{n-1} + 1^{n-1} + \ldots + (\mu_k - 1)^{n-1})) \pmod{p},$$

on applying (2). This gives

$$\frac{(mb + k)^n - m^n b_n}{n} \equiv m^n(0^{n-1} + 1^{n-1} + \ldots + (\mu_k - 1)^{n-1}),$$

modulo p. Setting $n = 1, 2, \ldots, p - 1$, in turn we obtain by adding, since

$$1 + (ma) + (ma)^2 + \ldots + (ma)^{p-2} \equiv \frac{(ma)^{p-1} - 1}{ma - 1} \pmod{p},$$

unless $ma - 1 \equiv 0 \pmod{p}$,

$$\sum_{i=1}^{p-1} \frac{(mb + k)^i - m^i b_i}{i} \equiv m \text{ or } 0 \pmod{p}, \tag{9}$$

according as $\mu_1 \geqq \mu_k$ or $\mu_1 < \mu_k$. In particular the right-hand member is m if $k = 1$. Using the known relations,

$$S_n(p) \equiv pb_n \pmod{p},$$

$0 < g < p$, the first congruence following involving p terms on the left, we have

$$g^{p-2}(1 + 1 + 1 + \ldots + 1) \equiv pb_0 g^{p-2} \pmod{p^2},$$
$$g^{p-3}(1 + 2 + \ldots + p - 1) \equiv pb_1 g^{p-3} \pmod{p^2},$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$1^{p-2} + \ldots + (p - 1)^{p-2} \equiv pb_{p-2} \pmod{p^2},$$

and we obtain by addition

$$pg^{p-2} + \sum_{\substack{a=1 \\ a \neq g}}^{p-1} \frac{a^{p-1} - g^{p-1}}{a - g} \equiv -p\frac{R_g}{g} \pmod{p^2},$$

employing (6), and if we write, with $(a, p) = 1$,

$$q(a) = \frac{a^{p-1} - 1}{p},$$

then the above gives

$$g^{p-2} + \sum_{\substack{a=1 \\ a \neq g}}^{p-1} \frac{q(a) - q(g)}{a - g} \equiv -\frac{R_g}{g} \pmod{p}.$$

Now modulo p,

$$g^{p-2} + \sum_{\substack{a=1 \\ a \neq g}}^{p-1} \frac{q(a)}{a-g} - \sum_{\substack{a=1 \\ a \neq g}}^{p-1} \frac{q(g)}{a-g} \equiv \sum_{\substack{a=1 \\ a \neq g}}^{p-1} \frac{q(a)}{a-g} - q(g)g^{p-2} + g^{p-2}$$

using

$$\sum_{\substack{a=1 \\ a \neq g}}^{p-1} \frac{1}{a-g} \equiv \sum_{a=1}^{p-1} (a-g)^{p-2} \equiv g^{p-2} \pmod{p},$$

hence

$$\frac{1-q(g)}{g} + \sum_{\substack{a=1 \\ a \neq g}}^{p-1} \frac{q(a)}{a-g} \equiv -\frac{R_g}{g} \pmod{p}, \tag{9a}$$

which for $g = 1$ gives

$$\sum_{a=2}^{p-1} \frac{q(a)}{a-1} \equiv -2 \pmod{p}.$$

We now obtain another formula related to (5b). Employing[1] (relation (4))

$$(b(m, k) + m)^{n+1} - b_{n+1}(m, k) = m(n+1)k^{p-2},$$

and setting $n = p - 2$, we have after employing (3a),

$$\sum_{i=0}^{p-2} (-1)^i b_i(m, k) m^{p-1-i} \equiv -mk^{p-2},$$

modulo p, where $(mb + k)^s = b_s(m, k)$.

We shall now modify some of the ideas in a certain proof[3] of (8) to obtain new types of congruences involving the Bernoulli numbers and Fermat's quotient. We have[4]

$$\frac{(1-m^n)b_n}{nm^{n-1}} \equiv \sum_{k=1}^{m-1} \sum_{i=1}^{[kp/m]} i^{n-1} \pmod{p}, \tag{10}$$

where $[x]$ is the greatest integer in x, $m > 1$, $n > 1$, $(m, p) = 1$. For $n = 1$ we shall now show that

$$(m-1)b_1 \equiv \sum_{k=1}^{m-1} \sum_{i=1}^{[kp/r]} i^{n-1} \pmod{p}. \tag{10a}$$

We have for $(k, m) = 1$, $(m, p) = 1$.

$$pk = r_k + m\left[\frac{pk}{m}\right]; \qquad 0 < r_k < m.$$

Set $k = 1, 2, 3, \ldots, m - 1$ in turn and add we have since $pk_1 \equiv pk_2 \pmod{m}$ only if $k_1 = k_2$ and therefore the r's range in some order over 1, 2, ..., $m - 1$,

$$m\sum_{k=1}^{m-1}\left[\frac{pk}{m}\right] + m\frac{m-1}{2} \equiv 0 \pmod{p},$$

or

$$\sum_{k=1}^{m-1}\left[\frac{pk}{m}\right] \equiv b_1(m-1) \pmod{p}.$$

Now take (10) with (10a), put $n = 1, 2, \ldots, p - 1$, in turn and add, after multiplying each relation in turn by g^{h-1-n}, we find

$$\frac{m^{p-1}-1}{m^{p-2}}b_{p-1} - \frac{(m-1)}{2} + \sum_{s=2}^{p-2} g^{p-1-s}\frac{1-m^s}{sm^{s-1}}b_s \equiv -g^{p-2}\,\alpha(g, m) \pmod{p}, \quad (11)$$

since in (10) and (10a) when we add we have on the right

$$\sum_{k=1}^{m-1}\sum_{i=1}^{[pk/m]}(g^{p-2} + ig^{p-3} + i^2g^{p-4} + \ldots + i^{p-2}),$$

or, if $0 < g < p$,

$$\sum_{k=1}^{m-1}\sum_{\substack{i=1\\ i\neq g}}^{[kp/m]}\frac{i^{p-1}-g^{p-1}}{i-g} + g^{p-2}(p-1)\alpha(g, m),$$

where g^{p-2} occurs $\alpha(g, m)$ times in the right-hand member of (10) and (10a). Now (11) is satisfied identically for $m = 1$. It may also be written

$$(m^{p-1}-1)b_{p-1} - \frac{m-1}{2m} + \sum_{s=2}^{p-2}\frac{m^{p-1-s}-1}{s}g^{p-1-s}b_s \equiv -\frac{g^{p-2}}{m}\alpha(g, m) \pmod{p} \quad (11a)$$

Set $g = 1$, $m = 1, 2, \ldots, p - 1$ and add, we have, since for $m < p$, that $\alpha(1, m) = m - 1$,

$$(S_{p-1}(p) - p + 1)b_{p-1} + \sum_{s=1}^{p-2}\frac{b_s}{s} \equiv 0 \pmod{p}, \quad (12)$$

since

$$S_a(p) \equiv 0 \pmod{p}$$

for $a < p - 1$. Now we can write

$$S_{p-1}(p) \equiv pb_{p-1} \pmod{p^2}. \quad (13)$$

Using (13) and $pb_{p-1} \equiv -1 \pmod{p}$, together with[5]

$$b_{p-1} \equiv 1 + \frac{(p-1)!}{p} \pmod{p},$$

we obtain the known[6] formula

$$\sum_{s=1}^{p-2} \frac{b_s}{s} \equiv W \pmod{p}, \tag{14}$$

where W is the Wilson quotient

$$\frac{(p-1)!+1}{p}.$$

On using this in connection with (11a) for $m < p$ we have another known[6] result

$$-q(m) + \sum_{s=1}^{p-2} \frac{b_s}{sm^s} \equiv W \pmod{p}. \tag{15}$$

Multiplying this last congruence through by a^t gives,[6] if t is in the range $1, 2, 3, \ldots, p-2$,

$$\sum_{a=1}^{p-1} a^t q(a) \equiv -\frac{b_t}{t} \pmod{p}. \tag{16}$$

The relation (15) is an analogous theorem to (6). If $m > 1$, $(m, p) = 1$, p prime and odd, then it is known that[7]

$$\frac{(m^n - 1)b_n}{n} \equiv \sum{}' (-1)^{n-1} \frac{f_{n-1}(\rho)}{1 - \rho^p} \pmod{p} \tag{17}$$

where

$$f_k(x) = 0^k + 1^k x + 2^k x^2 + \ldots + (p-1)^k x^{p-1}$$

with $0^0 = 1$. In this relation set $n = 2, 3, \ldots, p-1$, together with the relation

$$0 \equiv r^{p-2} \sum{}' \frac{\rho + \rho^2 + \ldots + \rho^{p-1}}{1 - \rho^p}$$

and add these $p - 1$ congruences, we obtain, if $r < p$, the symbol $\sum'$ indicating summation over all the distinct mth roots of unity excepting unity,

$$\sum_{s=1}^{p-1} \sum{}' \frac{\rho^s (r^{p-2} + r^{p-3} s + \ldots + s^{p-2})}{1 - \rho^p}$$

$$\equiv \frac{r^{p-2}(p-1)\rho^r}{1-\rho^p} + \sum{}' \sum_{\substack{s_1=1 \\ s_1 \neq r}}^{p-1} \frac{s_1^{p-1} - r^{p-1}}{s_1 - r} \div (1-\rho^p)$$

$$\equiv \sum{}' \frac{r^{p-2}\rho^r}{\rho^p - 1},$$

modulo p, so we obtain, if we set

$$\beta(m, i) = \sum{}' \frac{\rho^i}{(\rho^p - 1)},$$

for $m > 1$, $0 < r < p$,

$$-q(m) + \sum_{n=2}^{p-2} \frac{1 - m^n}{n} r^{p-1-n} b_n \equiv r^{p-2}\,\beta(m, r) \pmod{p}. \qquad (18)$$

But by (15) we also have

$$-\frac{r^{p-2}}{2} - q(r) + \sum_{n=2}^{p-2} \frac{r^{p-1-n} b_n}{n} \equiv W \pmod{p}. \qquad (19)$$

Now (18) may be written

$$-q(m) + \sum_{n=2}^{p-2} \frac{r^{p-1-n} b_n}{n} - \sum_{n=2}^{p-2} \frac{m^n r^{p-1-n} b_n}{n} \equiv r^{p-2}\beta(m, r),$$

modulo p, which with (19) gives, if $m > 1$, $0 < r < p$,

$$-q(m) + q(r) + \frac{r^{p-2}}{2} + W - \sum_{n=2}^{p-2} \frac{m^n r^{p-1-n} b_n}{n} \equiv r^{p-2}\beta(m, r), \quad (20)$$

modulo p. For $r = 1$ in particular this takes the form

$$-q(m) + W + \frac{1}{2} - \sum_{n=2}^{p-2} \frac{m^n b_n}{n} \equiv \beta(m, 1). \qquad (21)$$

Note that there is a certain type of analogy between (6b) and (20) with b_n corresponding to b_n/n.

[1] *Duke Math. Jour.*, **8**, 576 (1941).

[2] This relation is certainly one of the simplest the author has ever noted concerning the Bernoulli numbers. I have failed to find it in the literature, but do not imagine it could be new. Another proof may be obtained of it by the use of (2) by setting $n = p - 1$, and $k = p - 1$ therein.

[3] These PROCEEDINGS, **31**, 59, 312–314 (1945).

[4] *Annals Math.*, **18**, 114 (1917); *Duke Math. Jour.*, **3**, 571, relation (7) (1937).

[5] This relation is well known; for a recent proof, cf., Vandiver, *Duke Math. Jour.*, **8**, 578 (1941).

[6] Lerch, *Math. Annalen*, **60**, 471–490 (1905); Nielsen, *Oversigt Danske Vidensk. Selsk. Forhandlinger*, 518–519, 177–180 (1915).

[7] Included in a result due to Frobenius, *Berlin Sitsungsberichte*, 827 (1910). The latter theorem was generalized by Vandiver, *Trans. Amer. Math. Soc.*, **51**, 515 (1942).

CRITICAL POINTS OF HARMONIC FUNCTIONS AS POSITIONS OF EQUILIBRIUM IN A FIELD OF FORCE

BY J. L. WALSH

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated January 29, 1948

The first mechanical interpretation of the critical points (zeros of the derivative) of a polynomial $p(z)$ is due to Gauss,[1] who showed that the critical points not multiple zeros of $p(z)$ are the positions of equilibrium in the field of force due to unit particles situated at the zeros of $p(z)$, where each particle repels with a force equal to the mass divided by the distance. Bôcher (1904) showed that the zeros of the jacobian of two binary forms and of certain other algebraic invariants can be expressed as the positions of equilibrium in a similar field of force due to positive and negative particles situated at the zeros of the ground forms. Bôcher's results can be applied in the study of the critical points of rational functions.

A given harmonic function can frequently be expressed as the uniform limit of the logarithm of the modulus of a variable polynomial or more general rational function; the critical points of the harmonic function are the limits of those of the rational functions; the present writer (1933 and later) has thus applied the extensive theory of the location of the critical points of a rational function to the study of the critical points of a harmonic function. It is the object of the present note to show that the latter critical points can be conveniently and effectively studied *directly* as positions of equilibrium in a suitably chosen field of force, due to matter spread over Jordan curves or other point sets. The advantages of the present method are a closer analogy between the theory for harmonic functions and that for rational functions, with simpler proofs and stronger theorems, and in addition a systematic method for the enumeration of critical points of harmonic functions. We include also the means for studying various new harmonic functions.

A finite *critical point* of an analytic function $f(z)$ is a zero of the derivative $f'(z)$; the order of the critical point is the order of the zero of $f'(z)$. The point at infinity is a critical point of an analytic function $f(z)$, and of order k, if $z = 0$ is a critical point of $f(1/z)$, and of order k. If $f(z)$ has a pole at infinity, that point is not a critical point; if $f(z)$ is analytic there, the point at infinity is a critical point of order k if and only if $f'(z)$ has a zero there of order $k + 2$. A point $z = x + iy$ is a critical point of a harmonic function $u(x, y)$, and of order k, if it is a critical point of the analytic function $u + iv$, where v is conjugate to u, and of order k.

We shall consider critical points as positions of equilibrium in a field where the force is the conjugate of the derivative $f'(z)$ of an analytic func-

tion $f(z)$. Corresponding to the convention just introduced, we consider the point at infinity to be a position of equilibrium if and only if it is a critical point of $f(z)$.

1. We use the term *Jordan configuration* (of the extended plane) to denote either a continuum composed of a finite number of Jordan arcs or a finite number of such continua, mutually disjoint. Various integrals of potential theory representing a harmonic function $u(x, y)$ in a region R may be taken over the boundary B of R if B is a finite Jordan configuration; ordinarily such an integral is first taken over a level locus L with no multiple points, say $u(x, y) = \text{const}$ interior to R, and on L involves $(\partial u/\partial v)ds = -dv$, where $v(x, y)$ is conjugate (locally single valued) to $u(x, y)$ in R, and ν represents interior normal. If $u(x, y)$ is constant on a Jordan configuration J belonging to B, by allowing L to approach J the corresponding integral can perhaps be taken over J itself, on which $v(x, y)$ is continuous. Theorems 1, 5 and 11 below are all proved by this method from the formula

$$u(x, y) = \frac{1}{2\pi}\int_B \left(\log r \frac{\partial u}{\partial \nu} - u \frac{\partial \log r}{\partial \nu}\right) ds, \ (x, y) \text{ in } R, \qquad (1)$$

after making suitable simplifications.

THEOREM 1. *Let the boundary B of an infinite region R be a finite Jordan configuration, and let $G(x, y)$ be Green's function for R with pole at infinity. Then we have*

$$G(x, y) = \int_B \log r \, d\sigma + g, \qquad 0 \leqq \sigma \leqq 1,$$

where $d\sigma = -dH/2\pi$ on B, $\int_B d\sigma = 1$, $H(x, y)$ is conjugate to $G(x, y)$ in R, and g is a constant.

The function

$$H(x, y) = \int_B \arg (z - t)d\sigma(t), \qquad z = x + iy,$$

is conjugate to $G(x, y)$ in R, locally single valued, and the function

$$F(z) = G(x, y) + iH(x, y) = \int_B \log (z - t)d\sigma(t) + g$$

is analytic in R.

The critical points of $G(x, y)$ in R are the zeros in R of the analytic function

$$F'(z) = \int_B \frac{d\sigma(t)}{z - t}, \qquad z \text{ in } R,$$

and thus are the positions of equilibrium in the field of force due to the spread σ of positive matter on B, where the matter repels with a force proportional to the mass and inversely proportional to the distance.

The last part of Theorem 1 follows by taking the conjugate of $F'(z)$, for

the vector $1/(\bar{z} - \bar{t})$ represents the force at z due to a particle of unit mass at t.

Consideration of the direction of the field of force leads to an analog of Lucas' Theorem:

THEOREM 2. *Under the conditions of Theorem 1, all critical points of $G(x, y)$ in R lie in the smallest convex point set* Π *containing B. No boundary point z_0 of* Π *in R is a critical point of $G(x, y)$ unless B lies on a line through z_0.*

The latter part of Theorem 2 is difficult if not impossible to prove by previously published methods, as is the latter part of Theorem 3 below.

2. If in Theorem 1 the boundary B is symmetric in the axis of reals, the element of mass $d\sigma$ is equal at symmetric points of B. The integral expressing $F'(z)$ can be taken over one half of B, using a suitable convention for any part of B that may lie on the axis, with the integrand

$$\frac{1}{z - t} + \frac{1}{z - \bar{t}}. \tag{2}$$

If z is not real and lies exterior to the *Jensen circle* constructed on the segment $(t, \bar{t})$ as diameter, the conjugate of (2) is the force at z due to unit particles at t and $\bar{t}$, and whether or not t is real has a non-zero vertical component directed away from the axis; we have the first part of

THEOREM 3. *Under the conditions of Theorem 1, let B be symmetric in the axis of reals, and let Jensen circles be constructed with diameters the segments joining all pairs of symmetric points of B. Then all critical points of $G(x, y)$ in R lie in the closed interiors of these Jensen circles.*

A critical point z_0 in R not interior to a Jensen circle cannot lie on such a circle unless it lies on all Jensen circles for B, that is to say, unless B lies wholly on the equilateral hyperbola with vertices z_0 and $\bar{z}_0$.

The latter part of Theorem 3 follows from the fact that a point z on the Jensen circle for t and $\bar{t}$, the force represented by the conjugate of (2) is horizontal.

At any point $z_0 = x_0 + iy_0$ of R not a critical point, the force in the field described in Theorem 1 is orthogonal to the locus $G(x, y) = G(x_0, y_0)$; for if we compute dF/dz, where $dz = \epsilon\, d\nu (|\epsilon| = 1)$ is taken in the direction of the normal to this locus in the sense of increasing G, we have $dF/dz = (\partial G/\partial \nu)/\epsilon$, $\partial G/\partial \nu > 0$, $\partial H/\partial \nu = 0$, whence $\arg [F'(z)] = -\arg \epsilon$, so the direction of the force is that of ν; the conjugate of $F'(z)$ is the gradient of G. This remark obviously applies generally to the field defined by the conjugate of an analytic function $F'(z)$, where $F = G + iH$.

By study of the variation in the direction of the force as z traces a level locus near B, and also traces a suitably chosen auxiliary Jordan curve in R containing K, we study the variation of $\arg [F'(z)]$ on these curves, and obtain

THEOREM 4. *Under the conditions of Theorem 3 let K be a configuration*

consisting of a segment $\alpha \leqq z \leqq \beta$ plus the closed interiors of all Jensen circles intersecting that segment. Let α and β lie in R and not be critical points of $G(x, y)$. Let K contain k components of B. Then K contains $k - 1$, k, or $k + 1$ critical points of G in R according as the forces at α and β are both directed outward, one outward and one inward, or both directed inward.

3. For forces of the kind described, we prove the

LEMMA. *The force at a point P due to a distribution of positive matter in a circular region C not containing P is equal to the force at P due to the same total mass concentrated at a suitably chosen point of C.*

A *circular region* is the closed interior or exterior of a circle, or a half-plane.

This Lemma may be proved by inversion in P; the force at P due to a particle at Q is in magnitude, direction and sense the vector $Q'P$, where Q' is the inverse of Q; the proposition is equivalent to the fact that if a distribution of positive mass lies in the closed interior of a given circle, so does its center of gravity. This Lemma enables us to prove many further results, here omitted, from Theorem 1.

The method of proof of Theorem 1 also yields

THEOREM 5. *Let R be a region of the extended plane whose boundary B is finite and consists of disjoint Jordan configurations C and D. Let the function $u(x, y)$ be harmonic in R, continuous on $R + B$, zero on C and unity on D. Then in R we have*

$$u(x, y) = \int_C \log r \, d\sigma - \int_D \log r \, d\sigma + u_0, \qquad \int_C d\sigma = \int_D d\sigma, \quad (3)$$

$d\sigma = -dv/2\pi$ on C, $d\sigma = dv/2\pi$ on D, where u_0 is a constant and $v(x, y)$ is conjugate to $u(x, y)$ in R. We also have

$$v(x, y) = \int_C \arg (z - t) d\sigma - \int_D \arg (z - t) d\sigma,$$
$$f(z) = u + iv = \int_C \log (z - t) d\sigma - \int_D \log (z - t) d\sigma + u_0.$$

The finite critical points of $u(x, y)$ in R are the finite zeros in R of

$$f'(z) = \int_C \frac{d\sigma}{z - t} - \int_D \frac{d\sigma}{z - t}. \quad (4)$$

Consequently the critical points of $u(x, y)$ in R are precisely the positions of equilibrium in the field of force due to a spread σ of positive matter on C and a spread $-\sigma$ of negative matter on D, where the matter repels with a force proportional to the mass and inversely to the distance.

If R is finite the constant u_0 is zero or unity according as R is separated from the point at infinity by C or D; if R is infinite, $u_0 = u(\infty)$.

As an analog of a theorem due to Bôcher we prove

THEOREM 6. *Under the conditions of Theorem 5, let C and D lie, respectively, in disjoint circular regions S and T. Then all critical points of $u(x, y)$ in R lie in S and T. If C and D consist, respectively, of m and n*

components, then S and T contain, respectively, $m - 1$ and $n - 1$ critical points.

At a point z exterior to S and T the force due to the positive mass on C is equal to the force at z due to the same mass concentrated at some point Q_1 of S, by the Lemma, and the force at z due to the negative mass on D is equal to the force at z due to the same mass (numerically equal to the total positive mass) concentrated at a point Q_2 of T; the field due to these two masses at Q_1 and Q_2 has as lines of force the circles through Q_1 and Q_2, so the total force at z is not zero. Study of the variation in the direction of the force as z traces a circle separating S and T, and traces loci $u(x, y) = \epsilon$ and $1 - \epsilon$ in R, where ϵ is small but positive, completes the proof.

4. In Theorems 5 and 6 the constant values zero and unity on C and D may be replaced by any other distinct constant values; such a change may modify $u(x, y)$ by an additive constant, which alters u_0 in (3) but leaves (4) unchanged, and may modify $u(x, y)$ by a multiplicative constant, which multiplies u_0, $v(x, y)$, and hence $d\sigma$ by this same constant, and leaves (3) and (4) formally unchanged; even if the multiplicative constant is negative the conclusion of Theorem 5 persists. This remark will be useful in the proof of

THEOREM 7. *Let $u(x, y)$ be harmonic in a region R bounded by a Jordan curve J and a Jordan configuration B disjoint from J, continuous on $R + J + B$, zero on J and unity on B. Let the region R_1 bounded by J and containing R be provided with non-euclidean (NE) geometry by mapping R_1 conformally onto the interior of a circle, and let Π be the smallest NE convex set in R_1 containing B. Then all critical points of $u(x, y)$ in R lie in Π.*

We make the permissible assumption that J is the unit circle and that the origin is a point of R. We extend $u(x, y)$ harmonically across J so that $u(x, y)$ is harmonic in the region R' bounded by B and the reflection B' of B in J; then $u(x, y)$ is continuous on $R' + B + B'$, equal to unity on B and minus unity on B'. Both B and B' are finite. The values of $d\sigma$ are equal in corresponding points of B and B', so the integral corresponding to (4) for R' can be taken merely over B, with the integrand

$$\frac{1}{z - t} - \frac{1}{z - 1/\bar{t}}. \tag{5}$$

The conjugate of (5) is the force at z due to unit positive and negative masses at t and $1/\bar{t}$, and this force has the direction of the circle through z, t and $1/\bar{t}$, in the sense from t toward z. If z is a point of R exterior to Π, there exists a NE line L through z disjoint from Π, and each vector represented by the conjugate of (5) with t on B has a non-zero component orthogonal to L directed toward the side of L on which Π does not lie. Thus z is not a position of equilibrium and Theorem 7 is established. It

follows also that no boundary point of Π is a critical point unless Π reduces to a segment of a NE line.

If B in Theorem 7 is symmetric in a NE line L, there exist precise analogs of Theorems 3 and 4, where now the Jensen circles are to be taken as NE circles with NE centers on L. Still another related result is

THEOREM 8. *Let $u(x, y)$ be harmonic in a region R bounded by a Jordan curve J and two Jordan configurations B_1 and B_2 disjoint mutually and from J, continuous on $R + J + B_1 + B_2$, zero on J, unity on B_1, minus unity on B_2. Let the region R_1 bounded by J and containing R be provided with a NE geometry. If there exists a NE line L separating B_1 and B_2 in R, then the totality of such lines L separates a NE convex set Π_1 containing B_1 from a NE convex set Π_2 containing B_2; all critical points of $u(x, y)$ in R lie in Π_1 and Π_2.*

5. For a region R we denote generically by $G(z, z_0)$ Green's function with pole in z_0, and we study

$$u(z) = \sum_{k=1}^{m} \lambda_k G(z, \alpha_k), \qquad \lambda_k > 0. \tag{6}$$

THEOREM 9. *Let R be a Jordan region, and let the points $\alpha_1, \alpha_2, \ldots, \alpha_m$ of R lie in a smallest NE convex set Π. Then all critical points of $u(z)$ defined by (6) also lie in Π.*

Theorem 9 can be proved by remarking that for large positive M the level locus $u(z) = M$ in R consists of m curves, one about each point α_k; Theorem 7 then applies; as M becomes infinite the smallest NE convex set containing the locus approaches Π.

In a similar manner, by considering the level loci $u(z) = M$ and $u(z) = -M$, there may be proved

THEOREM 10. *Let R be a Jordan region, and let $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ be points of R which lie in two NE convex sets Π_1 and Π_2 separated by every NE line separating the α_j from the β_k. Then all critical points of*

$$u(z) = \sum_{k=1}^{m} \lambda_k G(z, \alpha_k) - \sum_{k=1}^{n} \mu_k G(z, \beta_k), \qquad \lambda_k > 0, \mu_k > 0, \tag{7}$$

in R lie in Π_1 and Π_2.

The situation of Theorem 9 can be studied directly even under a more general hypothesis. In Theorem 1 let z_0 be a finite point of R; choose $u(x, y)$ in (3) as $G(z, z_0)$, with C as B and D as the locus $G(z, z_0) = M$. When M varies, $v(x, y)$ does not change, and we have $\int_D d\sigma = 1$, so as M becomes infinite we obtain

$$G(z, z_0) = \int_B \log r \, d\sigma - \log |z - z_0| + u_0, \qquad \int d\sigma = 1,$$
$$F(z) = G(z, z_0) + iH(z, z_0) = \int \log (z - t) d\sigma - \log (z - z_0) + u_0,$$

$$F'(z) = \int_B \frac{d\sigma}{z-t} - \frac{1}{z-z_0}, \qquad \int_B d\sigma = 1.$$

THEOREM 11. *Let R be a region bounded by a finite Jordan configuration B, and let $\alpha_1, \ldots, \alpha_m$ be finite points of R. Then the critical points of $u(z)$ defined by (6) are the positions of equilibrium in a field of force due to a spread σ of positive matter of total mass $\Sigma\lambda_k$ on B, and to negative particles of respective masses $-\lambda_k$ at the α_k.*

One of many consequences of Theorem 11 is

THEOREM 12. *Under the conditions of Theorem 11 let B lie in a circular region S and the α_k in a circular region T disjoint from S. Then all critical points of $u(z)$ in R lie in S and T.*

6. As is well known, equation (1) expresses $u(x, y)$ as the sum of a simple potential and a double layer potential; in any case we may write from (1) at least formally[2]

$$u(x, y) = \frac{-1}{2\pi}\int_B \log r \, dv + \frac{1}{2\pi}\int_B u \, d[\arg (z-t)]$$

$$= \frac{-1}{2\pi}\int_B \log r \, dv + \frac{1}{2\pi}[u \arg (z-t)]_B - \frac{1}{2\pi}\int \arg (z-t) du,$$

$$f(z) = u + iv = \frac{-1}{2\pi}\int_B \log (z-t) dv - \frac{i}{2\pi}[u \log (z-t)]_B + \frac{i}{2\pi}\int_B \log (z-t) du,$$

$$f'(z) = \int_B \frac{d\sigma}{z-t} - \frac{i}{2\pi}\left[\frac{u}{z-t}\right]_B + \frac{i}{2\pi}\int_B \frac{du}{z-t}. \tag{8}$$

The gradient of $u(x, y)$ is the conjugate of $f'(z)$, for which there is an obvious interpretation as a field of force due to spreads σ and iu of matter over B and particles at the end-points of B. Theorems 1, 5 and 11 reduce this general configuration to simpler manageable situations for suitable $u(x, y)$, and those theorems by no means exhaust the possibilities. For instance, if $u(x, y)$ has the constant value unity on an arc A: (α, β) of B, the corresponding last two terms of (8) reduce to

$$\frac{-i}{2\pi}\left[\frac{1}{z-\beta} - \frac{1}{z-\alpha}\right]; \tag{9}$$

the conjugate of (9) is a field of force due to *skew particles* at β and α, of masses $+1$ and -1, respectively; these particles exert forces at z equal to the inverse distance, in the directions $\arg (z-\beta) + \pi/2$ and $\arg (z-\alpha) - \pi/2$. The lines of force in this field are the circles of the coaxial family determined by α and β as null circles. Such fields of force may be

combined with similar ones, or with such fields as occur in Theorems 1, 5 and 11.

Moreover, if $u(x, y)$ is harmonic interior to a circle C and is extended harmonically across an arc of C on which $u(x, y)$ vanishes, $\partial u/\partial \nu$ has equal and opposite values on the two sides of the complementary arc of C, so over the latter arc the first part of the integral in (1) or (8) vanishes. Several results are consequences of these remarks:

THEOREM 13. *Let the points* $\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_n = \alpha_0, \beta_n = \beta_0$ *lie on the unit circle* C: $z = e^{i\theta}$ *in the counterclockwise order indicated, and let the function* $u(z)$ *harmonic and bounded interior to* C *take a positive constant value* γ_k *on each arc* $\alpha_k < \theta < \beta_k$ *and the value zero on each arc* $\beta_k < \theta < \alpha_{k+1}$. *Let NE lines* $A_0, A_1, \ldots, A_{n-1}$ *be so chosen that* A_k *is an arc of a circle belonging both to the coaxial family determined by* α_k *and* β_k *and to that determined by* α_{k+1} *and* β_{k+1}. *Then the NE convex set interior to* C *bounded by the* A_k *contains all critical points of* $u(z)$.

These arcs A_k are special cases of W-curves, which are of importance in the entire study of critical points, namely, the locus of points at which the forces due to two groups of particles have the same direction but opposite senses; the groups of particles (here the n groups α_k, β_k) have relations of equality of mass between the elements of a group, but between groups all possible relative masses of constant sign are considered. As in Theorem 13, the W-curves form the boundaries of regions necessarily containing the critical points. A double layer distribution on an arbitrary curve C is the limit of a double layer distribution piecewise constant on C as in Theorem 13; the lines of force for the original distribution are the circles normal to C, and the W-curves are the circles normal to C in two points plus the locus of points z_0 from which two circles tangent at z_0 and normal to C can be drawn.

THEOREM 14. *Let* R *be a region bounded by the unit circle* C *and by a Jordan configuration* B *interior to* C. *Let the function* $u(x, y)$ *be harmonic and bounded in* R, *continuous on* $R + B + C$ *except in the end-points of an arc* A *of* C, *unity on* B *and the interior points of* A, *zero on the interior points of* $C - A$. *Let* Π *be a NE convex set containing* B, *and let the set* Π_1 *consist of* Π *plus all points of arcs joining* A *and* Π_1 *of circles of the coaxial family determined by the end-points of* A *as null circles. Then all critical points of* $u(x, y)$ *in* R *lie in* Π_1.

The conclusion of Theorem 14 applies also to an arbitrary linear combination with positive coefficients of the two harmonic measures $\omega(z, A, |z| < 1)$ and $\omega(z, B, R)$.

Theorem 14 is a sharpened analog of Theorem 7; a similar sharpened analog of Theorem 8 also exists, and as in Theorem 13 we may replace a single arc of C by a sum of arcs of C; both positive and negative values of $u(x, y)$ on B and C may be admitted. Indeed, all the results formulated

in the present note are intended only as illustrations of an extensive theory which the writer plans to develop elsewhere.

[1] For detailed references to the literature, the reader may consult a forthcoming book: Marden, M., *Geometry of the Zeros of a Polynomial*, to be published as a volume of Mathematical Surveys by the American Mathematical Society.

[2] Equation (8) is Cauchy's integral formula for $f'(z)$, involving an unusual term to include discontinuous $u(x, y)$.

SOME PROPERTIES OF ROTATIONAL FLOW OF A PERFECT GAS

BY P. NEMÉNYI AND R. PRIM

NAVAL ORDNANCE LABORATORY, WHITE OAK, MARYLAND

Communicated by S. Lefschetz, January 14, 1948

1. Introduction.—Complete analytic solutions of definite boundary value problems of compressible rotational flow are extremely difficult to obtain. Therefore, in the present stage of study of rotational gas flow the inverse, or semi-inverse, approach promises to prove fruitful. We propose to deduce and to study flow patterns satisfying the differential equations of gas flow and having throughout the field certain prescribed geometric, kinematic or physical properties; but will not make them obey any prescribed boundary conditions.

Usually entire systems of flow patterns, rather than individual solutions, result from this type of approach. If an imposed condition is proved not to be satisfied by any possible flow pattern, then this negative result is of significance, as it establishes a general property of all flows satisfying the equations of the problem.

In various fields of fluid mechanics the semi-inverse approach has been used with considerable success. It is sufficient to recall the contributions of Beltrami, Massotti, Jeffreys, Hamel, Oseen, Kampé de Fériet,[1] Bateman and Tollmien.[2] None of these investigations deals however with the rotational flow of a perfect gas.

The reason for this is probably that such an investigation—unless restricted to cases of uniform stagnation enthalpy—requires complicated and difficult eliminations if based on the familiar equations of the problem.

In the recent investigations of Munk and Prim[3, 4] this elimination is accomplished with complete generality as far as steady flow of a perfect gas is concerned. This makes the application of the semi-inverse approach to rotational gas flow practicable. Therefore the results of Munk and Prim are used throughout the present paper. The present paper is

a general preliminary report on our investigations, parts of which will be published in detail elsewhere.

2. Some Properties of the Plane Flow of a Perfect Gas.—A natural coördinate system for the analysis of problems of plane fluid motion employs as coördinate curves the streamlines (ξ = const) and their orthogonal trajectories (η = const). In this coördinate system the vorticity of the reduced velocity field is expressed by

$$|\operatorname{curl} \bar{w}| = -\frac{1}{g_1 g_2} \frac{\partial}{\partial \xi} g_2 w \tag{1}$$

($g_1\, d\xi$ and $g_2\, d\eta$ are components of the vector element of arc length).

The vorticity equation and the continuity equation[4] take on the following form:

$$\frac{\partial}{\partial \eta}\left[\frac{w^2}{1-w^2} \frac{\partial \ln (g_2 w)}{\partial \xi}\right] = 0 \tag{2}$$

$$\frac{\partial}{\partial \eta} [g_1 (1-w^2)^{\beta} w] = 0. \tag{3}$$

These equations enable us to settle certain general questions concerning plane gas flows.

We ask first under what conditions the velocity is constant along each streamline. Since the ultimate velocity is in every case constant along each streamline, our question is equivalent to asking under what conditions the reduced velocity is constant along each streamline.

Introduction of the requirement $w = w(\xi)$ into equations (2) and (3) yields the geometric conditions:

$$\frac{\partial^2 \ln g_2}{\partial \xi \partial \eta} = 0 \qquad \text{or } g_2 = a(\xi) b(\eta)$$

$$\frac{\partial \ln g_1}{\partial \eta} = 0 \qquad \text{or } g_1 = c(\xi).$$

An analysis of this geometric limitation leads to the following theorem:

THEOREM 1. *In plane steady flow of a perfect gas, in absence of an external field of mass force, the velocity magnitude can have a constant value along each streamline only if the streamlines are concentric circles or parallel straight lines.*

Obviously, in these cases in which w = const (and hence also v = const) along each streamline, curl $\bar{w}$ and curl $\bar{v}$ also have a constant magnitude along the same lines. The question arises whether the converse theorem is also true, that is, whether the only two dimensional flow patterns possible in a perfect gas such that the streamlines are lines of equal vorticity are those in which they are also lines of equal velocity magnitude.

To settle this question we consider first flow with constant reduced vorticity magnitude along each streamline. Crocco's pressure theorem as reformulated[5] for a general distribution of stagnation enthalpy requires that the pressure remain constant along each streamline if the reduced vorticity does so; that is, $p = p(\xi)$. On the other hand, according to a familiar relation

$$\frac{p}{p_0(\xi)} = (1 - w^2)^{\beta+1}$$

(p_0 denotes the stagnation pressure, which remains constant along each streamline between shock fronts). Hence $|\text{curl } \bar{w}| = f(\xi)$ implies that w must be also a function of ξ alone.

A very similar argument shows that if the lines of equal velocity magnitude (isovels) and the lines of equal vorticity (isocurls) of the reduced velocity field coincide, then they coincide also with the streamlines.

Starting from the vector identity

$$\text{curl } \bar{v} = a(\xi) \text{ curl } \bar{w} + \text{grad } a(\xi) \times \bar{w}$$

we have transformed the above statements relating to the reduced velocity field into analogous statements concerning the actual velocity field and proved the following theorems:

THEOREM 2. *In plane steady flow of a perfect gas, in absence of an external field of mass forces, constant vorticity along each streamline implies constant velocity along the same lines.*

THEOREM 3. *In plane steady flow of a perfect gas, in absence of an external field of mass forces, the coincidence of the isovels and the isocurls implies that they coincide also with the streamlines.*

Our three theorems prove that all flow fields having coincident isovels and streamlines, or coincident isocurls and streamlines, or coincident isovels and isocurls are restricted to flows the streamlines of which are concentric circles or parallel straight lines.

In a paper by R. Prim[6] the question is investigated whether the streamline pattern can form, with its orthogonal trajectories, an isothermal net, that is a net in which $g_1 = g_2$. The following theorem is established.

THEOREM 4. *In plane steady flow of a perfect gas in absence of an external field of mass forces, the streamlines and their orthogonal trajectories can form an isothermal net only if the streamlines are concentric circles, parallel straight lines or radial straight lines: the last case being possible only if the reduced velocity field is irrotational.*

It will be instructive to compare the geometric restrictions imposed by the above four conditions upon the pattern of gas flow with the corresponding restrictions imposed upon incompressible non-viscous and in-

compressible viscous flow. (For all three types of idealized fluids absence of an external field of mass forces is assumed.)

	NON-VISCOUS INCOMPRESSIBLE FLUID ("IDEAL FLUID")	VISCOUS INCOMPRESSIBLE FLUID	PERFECT GAS
Streamlines coincide with isovels	Streamlines concentric circles or parallel straight lines	Answer is not known	Streamlines concentric circles or parallel straight lines
Isocurls coincide with streamlines	*Any* rotational plane flow has this property (Helmholtz)	(1) Any flow which has a constant curl throughout the field (2) Streamlines concentric circles or parallel straight lines (Kampé de Fériet)	Streamlines concentric circles or parallel straight lines
Isocurls coincide with isovels	Streamlines concentric circles or parallel straight lines	Answer is not known	Streamlines concentric circles or parallel straight lines
Streamlines and their orthogonal trajectories form isothermal net ($g_1 = g_2$)	(1) Any irrotational flow (2) Rotational flow only if streamlines concentric circles or parallel straight lines	(1) Any irrotational flow (2) Rotational flow only if streamlines logarithmic spirals or their limiting cases (concentric circles, radial straight lines, parallel straight lines (Jeffreys, Hamel)	(1) Irrotational flow: streamlines concentric circles parallel straight lines or radial straight lines (2) Rotational flow: streamlines concentric circles or parallel straight lines

3. On Flow Fields for Which, in a Cylindrical Isothermal Coördinate System All Three Velocity Components Depend Only on One Coördinate.—For irrotational, plane gas flow this question was investigated by Tollmien.[2] We have extended the study to rotational flow, at the same time allowing an additional velocity component parallel to the elements of the cylindric coördinate surfaces. In one respect, however, our investigation is less general than that of Tollmien: we have investigated only systems of such cylindrical coördinate surfaces which in the xy plane form an *isothermal* net, while Tollmien admitted orthogonal nets of any kind.

The main result of the investigation can be summarized in the theorem:

THEOREM 5. *The only isothermal cylindrical system of coordinates for which the reduced velocity components of any steady non-parallel flow of a perfect gas, in absence of an external field of mass forces, are functions of one of the isothermal coordinates only are formed by logarithmic spirals and their limiting cases.* Thus the coördinate systems satisfying our kinematical

conditions are exactly the same as those found by Tollmien for plane potential flows. With reference to this spiral-cylindric coördinate system the differential equations can be reduced to a pair of ordinary non-linear differential equations of the first order.

4. Massotti Flow.—Massotti sought incompressible, non-viscous flow patterns for which $v_z = 0$, and for which at the same time $(\text{curl } \bar{v})_z = 0$.

There is a wide variety of compressible flow patterns the reduced velocity vector of which satisfies the corresponding relations $w_z = 0$ and $(\text{curl } \bar{w})_z = 0$. We satisfy the latter equation by setting

$$w_x = \frac{\partial \phi}{\partial x}; \qquad w_y = \frac{\partial \phi}{\partial y}.$$

Hence

$$w^2 = \left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2$$

and the continuity equation takes on the form:

$$\Delta_{x,y}\phi[1 - w^2] = \beta\left[\frac{\partial \phi}{\partial x}\frac{\partial w^2}{\partial x} + \frac{\partial \phi}{\partial y}\frac{\partial w^2}{\partial y}\right].$$

Our vorticity equation, $\text{curl}\left[\dfrac{\bar{w} \times \text{curl } \bar{w}}{1 - w^2}\right] = 0$, is automatically satisfied as $\bar{w} \times \text{curl } \bar{w}$ is obviously perpendicular to the x, y plane and equal to $^1/_2(\partial w^2/\partial z)$.

Hence the only condition imposed upon ϕ is the continuity equation which is identical with that of irrotational plane compressible flow. We can immediately see that a wide variety of Massotti flows can be built from families of plane irrotational flows in the following way. Assume $\phi(x, y, c_1, c_2, \ldots c_n)$ is the potential of the reduced field of an n-parameter family of irrotational flows. Replace each of these parameters by an arbitrary function of z; the ϕ thus obtained can be considered as the potential from which the x and y components of families of reduced velocity fields of the Massotti type can be derived.

For example, we have based upon the simple irrotational Prandtl-Meyer flow a pseudo-plane rotational flow past a developable surface.

5. Beltrami Flows and Other Investigations.—Beltrami[7] was the first to investigate velocity fields which satisfy the condition $\bar{v} \times \text{curl } \bar{v} = 0$. In recent memoranda[8, 9, 10] the present writers investigate flows of perfect gases which satisfy this relation (Beltrami Flow Proper). They found it fruitful to study also the broader family of flows for which $\bar{w} \times \text{curl } \bar{w} = 0$ (Generalized Beltrami Flow).

Further semi-inverse investigations under way include the study of gas

flow patterns for which all streamlines are straight lines or for which the hodograph space degenerates into a surface or a curve. In addition, the basis of these semi-inverse methods, the equations of Munk and Prim, are being generalized to include certain classes of non-steady flow and of compressible substances with a more general state equation.

[1] See, for example, Berker, A. Ratib, "Sur quelques cas d'integration des équation du mouvement d'un fluide visqueux incompressible," Paris et Lille, 1936.

[2] See Tollmien, W., "Zum Übergang von Unterschall—in Überschallströmungen," *Z. f. Angewandte Mathematik und Mechanik*, **17**, 117–136 (1937).

[3] Munk, M. M., and Prim, R., "On the Multiplicity of Steady Gas Flows Having the Same Streamline Pattern," *Proc. Nat. Acad. Sci.*, **33**, 137–141 (1947).

[4] Munk, M. M., and Prim, R., "On the Canonical Form of the Equations of Steady Motion of a Perfect Gas," Naval Ordnance Laboratory Memorandum No. 9169 (June, 1947).

[5] Prim, R., "Extension of Crocco's Theorems to Flows Having Non-Uniform Stagnation Enthalpy," Naval Ordnance Laboratory Memorandum No. 9286 (August, 1947); *Phys. Rev.*, Jan. 15, 1948.

[6] Prim, R., "On the Existence of Steady Gas Flow in Plane Isothermal Streamline Patterns," Naval Ordnance Laboratory Memorandum No. 9267 (August, 1947); *Bull. Amer. Math. Soc.* (forthcoming).

[7] Beltrami, E., "Considerazioni idrodinamiche," *Rendiconti Istituto Lombardo*, **22**, 121–130 (1889).

[8] Neményi, P., and Prim, R., "Some Patterns of Vorticose Flow of a Perfect Gas," Naval Ordnance Laboratory Memorandum No. 9219 (June, 1947).

[9] Neményi, P., and Prim, R., "Some General Properties of Beltrami Flows of a Perfect Gas," Naval Ordnance Laboratory Memorandum No. 9397 (October, 1947).

[10] Neményi, P., and Prim, R., "Some Interesting Cases of Steady Helicoidal Beltrami Flow of a Perfect Gas," Naval Ordnance Laboratory Memorandum No. 9426 (November, 1947).

REPRODUCTIVE DIAPAUSE IN DROSOPHILA ROBUSTA

BY HAMPTON L. CARSON AND HARRISON D. STALKER

DEPARTMENT OF ZOOLOGY, WASHINGTON UNIVERSITY, ST. LOUIS, MO.

Communicated by Th. Dobzhansky, January 19, 1948

Introduction.—Studies of microevolutionary processes in natural populations of *Drosophila* have been hampered by a lack of ecological data. For most species of the genus, the basic facts concerning population size, migration, longevity, specific breeding sites, length of breeding season and nature of overwintering populations are unknown. Workers in this field have not ignored these subjects[1, 2, 3], but it is clear that preoccupation with other basic problems has generally held up experiments specifically planned to answer these questions. It is, nevertheless, true that a full understanding of natural selection, genetic drift and other processes, as they are oper-

ative in natural populations of *Drosophila*, will not be attained until the ecology of the various species concerned is better known. The data presented in this paper are based largely on observations which were incidental to a study of the gene arrangements and morphology of *Drosophila robusta*, throughout the warm seasons of the year in a single locality near St. Louis, Missouri.

Although of a preliminary nature, these data show that a striking reproductive arrest, or diapause, sets in during the late summer and early autumn in the natural population, despite the fact that temperatures apparently favorable for the breeding of the flies in nature continue ordinarily until the middle of October. The importance of this fact for the interpretation of population dynamics in this species is obviously great, and further, more detailed investigations of this and similar subjects are planned.

Observations.—The specimens of *Drosophila robusta* on which the following observations were made were collected from paper cups baited with ripe banana mash and set in a particular deciduous woods, supporting considerable undergrowth, near Olivette, St. Louis County, Missouri. In 1946, flies were collected over nine periods of variable length between March 5 and December 10 (see first column of table 1). As the offspring of these

TABLE 1

FERTILITY OF MATURE WILD-CAUGHT FEMALES OF *Drosophila robusta*. 1946 COLLECTION

COLLECTION PERIOD	NO. OF MATURE* FEMALES COLLECTED	FEMALES PRODUCING OFFSPRING WITHIN SIX DAYS	
		NO.	PER CENT
I. Mar. 5–Apr. 6	38	27	71.1
II. Apr. 13–May 15	99	69	69.7
III. May 18–Jun. 1	153	120	78.4
IV. Jun. 17–Jun. 19	203	171	84.2
V. Jul. 3–Jul. 9	108	84	77.8
VI. Jul. 23–Aug. 6	75	64	85.3
VII. Aug. 15–Sept. 3	42	29	69.0
VIII. Sept. 13–Sept. 24	52	2	3.9
IX. Oct. 14–Dec. 10	45	0	0.0

* Wild-caught females which had obviously only recently emerged from the pupal case are not included in this table.

wild flies were to be used for the determination of the frequencies of the various gene arrangements and an F_2 generation from the majority of the wild flies obtained, the following procedure was adopted. Each wild-caught female was placed individually into a separate culture vial soon after capture; after three days it was transferred to a fresh culture vial and at the end of six days both vials were examined for the presence of larvae. If larvae were present, the female was recorded as having been inseminated in nature and the larvae were raised, some being used for the cytological determination of the salivary gland chromosome sequences and the re-

mainder saved as a mass F_1 from the original wild female. If no larvae were observed in six days, the female was mated to a male captured on or about the same date and the F_1 handled in a similar manner. From the middle of June on, a considerable number of females were captured which, because of their light coloration, had obviously only recently emerged from the pupal case. These were separated from the "mature" ones, aged for ten days en masse and then mated individually to males captured on the same day.

Table 1 shows the total number of females, judged as relatively mature, which were caught during the nine collection periods and the per cent of these which were found to have been inseminated in nature, as indicated by the production of offspring in the laboratory within six days. It will be seen that this percentage remains high from March 5 through August 6, drops in the next period (August 15 through September 3) and falls abruptly to a very low figure for females collected after September 13. It is probable, from the 1947 data given below, that most of the flies failing to produce offspring within six days were virgin, although from the 1946 data alone, the possibility of some sort of slowing-down of egg production by inseminated females cannot be excluded. With a few exceptions, however, all the females collected on or after September 13 produced offspring after being mated to males captured at the same time, although in many cases the production of fertile eggs by these females was delayed for a week or more after being placed with a male.

TABLE 2

ELAPSED TIME BETWEEN EMERGENCE AND THE FIRST APPEARANCE OF F_2 LARVAE FOR F_1 FLIES FROM WILD-CAUGHT *Drosophila robusta*. 1946 COLLECTION

COLLECTION PERIOD OF P_1 WILD FLIES	NO. OF F_1S OBTAINED	NO. OF DAYS ELAPSING FROM EMERGENCE OF F_1 FLIES UNTIL F_2 LARVAE WERE PRODUCED		
		MEAN	MODE	RANGE
IV. Jun. 17–Jun. 19	146	14.5	13	7–31
V. Jul. 3–Jul. 9	150	17.2	14	8–50
VI. Jul. 23–Aug. 6	106	19.5	15	7–51
VII. Aug. 15–Sept. 3	94	25.4	21	10–56
VIII. Sept. 13–Sept. 24	109	33.5	27	11–65
IX. Oct. 14–Dec. 10	31	33.0	34	13–55

F_2 generations were obtained in the following manner from most of the above wild-caught females or pairs of wild flies. F_1 flies of both sexes were maintained in mass culture in vials and changed every three days until F_2 larvae were observed. For each of these F_1s, the date of the latter event, as well as that of the emergence of F_1 flies was recorded as a matter of routine. These data are thus useful for an analysis of the speed with which the F_1 flies reached a condition of active reproduction as indicated by the production of offspring. Table 2 shows that the F_1s from wild flies caught from August 15 on attained this condition much more slowly than those

collected in June, July and the first week in August. Correspondence between these data and those for the fertility of the wild-caught parental flies (table 1) is striking, and it is clear that the reproductive delay manifested by the parental flies caught on or after August 15 is reflected in their offspring, despite the fact that the latter were kept under uniform laboratory conditions of temperature (25.5°C.).

The above data appear to be significant in view of the fact that in old laboratory stocks of *Drosophila robusta*, sexual maturity (as indicated by willingness to mate) is attained about four days after emergence for females and about eight days for males. In this species, mating appears to occur several days before the females produce offspring, although further study of this characteristic in laboratory stocks is desirable.

TABLE 3

INSEMINATION OF WILD FEMALE *Drosophila robusta* CAPTURED DURING SEPTEMBER AND OCTOBER, 1947

DATE OF CAPTURE	TOTAL NO. OF FEMALES	NO. OF FEMALES INSEMINATED IN NATURE
1. Insemination determined by observation of hatching eggs:		
Sept. 28–Oct. 2	14	0
Oct. 3–5	30	8
Oct. 7–9	27	1
Oct. 10–12	49	1
Oct. 13–15	28	0
2. Insemination determined by dissection:		
Oct. 14–15	18	1
Oct. 16–18	25	1
Oct. 20–22	21	1
Oct. 23–24	17	0
Oct. 25–27	26	0
Total	256	13
Per cent inseminated in nature		5.1

In 1947, collections were made during two periods from the same locality in the same woods. These periods extended from May 19 through June 7 and from September 28 through October 27, respectively. In order to reduce the labor involved and to expedite analysis of the chromosomes, each wild-caught female in the spring collection was placed with a male immediately upon capture. For this reason, the spring data are of no use in the analysis of fertility in nature. All but very few pairs produced offspring in six days; no F_2 generations were raised in 1947.

The fall (1947) collection was analyzed as follows. Females were placed individually into culture vials shortly after capture. At three-day intervals, each was changed to a fresh vial and the surface of the food closely examined under the binocular microscope for the presence of eggs or larvae. The date of the appearance of eggs was recorded, and if no larvae hatched

from these in the six days following, the female was considered to be virgin and was mated to a wild-caught male. In addition to these, 107 females caught between October 14 and October 27 (when the collecting season was ended by the abrupt onset of cold weather) were dissected immediately on being brought to the laboratory. For each of these, notes were made on the size of the ovary, the character and amount of body fat and whether the female had been inseminated in nature, as indicated by the presence of sperm in the ventral receptacles or spermathecae.

Table 3 shows that approximately 95% of the females caught during this period were virgin, as determined either by watching the fate of eggs laid in the laboratory (upper part of table) or by dissection (lower part of table). Of the 107 females which were dissected, furthermore, none had functional ovaries as indicated by the presence of large filamented ovarian eggs, although in two of them the ovaries were medium in size and contained filamented eggs approximately two-thirds mature size. In almost all the females, the ovaries were very small and rounded (approximately 380 μ in length), similar to those of newly emerged flies. A high percentage of the flies, furthermore, had heavy sheets of body fat, sometimes almost completely filling the abdomen. Such a condition is quite unlike that in a newly emerged fly in which the residual pupal fat, in the form of spherical "nodules" is easily recognized. In 13 of the females, nodular fat was found; this was taken as indicative of relative youth, although, in some of these, fat was also present in small sheets.

Of particular interest were the three dissected flies which had been inseminated in nature. Like most of the virgins, these flies had abundant body fat present in sheets and very small ovaries. The latter, however, were peculiarly elongated and flattened apically, while proximally they presented a slightly swollen appearance. This swollen area contained small ovarian ova of a distinct yellowish-brown color, quite unlike the colorless or whitish ova usually encountered. Ovaries of this unusual type were also found in one of the virgin flies.

Of the 148 females collected between September 28 and October 15, 10 had been inseminated in nature and 138 were virgin. The elapsed time between capture and the laying of the first infertile eggs by the virgins was recorded. This was extremely variable, ranging from 2–25 days, the mean number of days elapsing being 10.3. The 10 inseminated females, furthermore, required a similar mean number of days (9.1) before fertile eggs were laid. One of these females showed an elapsed time of 25 days from the time of capture until the first eggs were laid.

Conclusions.—The 1946 data show clearly that, beginning as early as August, wild females of *Drosophila robusta* captured in the fall reproduce much more slowly than those collected in spring and early summer. A similar delay is manifested in the offspring of these flies. The 1947 data

support this conclusion and suggest that the fall population of females, at least, consists principally of virgin flies with undeveloped ovaries and large quantities of body fat. At least some of the inseminated females have very small, peculiar ovaries which suggest regression from a former more highly developed and possibly functional condition. This relationship between the size of the ovary and the amount of body fat closely resembles the condition in the reproductive diapause which has been shown to occur in other insects in the fall (e.g., "dissociation gono-trophique" in *Anopheles maculipennis*[4]).

A somewhat similar diapause has been described for *Drosophila nitens*.[5] Laboratory strains of this species, despite the fact that they were kept at 25°C., suspended reproduction from November to March for two consecutive years. Bertani[6] has shown that this reproductive diapause can be broken by subjecting the adult flies to temperatures from 2–5°C. for 10 days. In *Drosophila robusta*, on the other hand, no interruption or delay of breeding in old laboratory stocks has been observed and a mean of ten days of laboratory conditions appears to be sufficient to break the diapause in wild-caught females.

Biologically, the inference to be drawn is that *Drosophila robusta* overwinters as an adult and that this hibernation period is preceded by a physiological change-over from egg production to the deposition of body fat. The factors, whether genetic or environmental or both, which may be responsible for the initiation of the diapause are unknown and will be the object of further investigation. From the point of view of population dynamics, however, the central fact which emerges from these observations is that the fall population of this species consists largely of non-breeding individuals.

[1] Dobzhansky, Th., and Epling, C., *Carnegie Inst. Wash. Publ. No.* **554**, pp. 1–46 (1944).

[2] Dobzhansky, Th., and Wright, S., *Genetics*, **28**, 304–340 (1943).

[3] Spencer, W. P., *Ohio Jour. Sci.*, **41**, 190–200 (1941).

[4] Swellengrebel, N. H., *Ann. Inst. Pasteur*, **43**, 1370–1389 (1929).

[5] Buzzati-Traverso, A., *Istituto Lombardo di Science e Lettere, Rendiconti*, **77**, 37–49 (1944).

[6] Bertani, G., *Nature*, **159**, 309 (1947).